This report is based on studies performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. The work was sponsored by the Naval Submarine Medical Research Laboratory under Air Force Contract F19628-90-C-0002.

The ESC Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Gary Tutungian
Administrative Contracting Officer
Directorate of Contracted Support Management

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

SIGNAL ENHANCEMENT IN AM-FM INTERFERENCE

T.F. QUATIERI
R.B. DUNN
R.J. McAULAY
Group 24

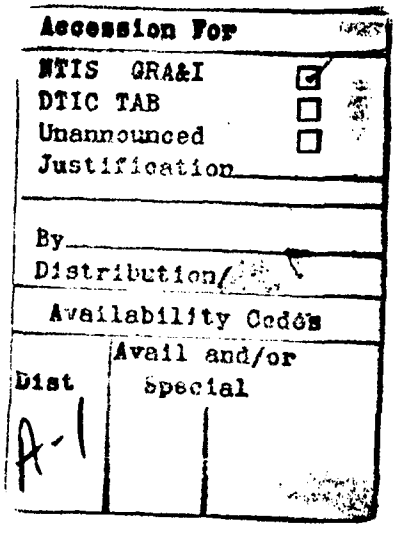
TECHNICAL REPORT 993

17 MAY 1994

LEXINGTON MASSACHUSETTS

ABSTRACT

A new approach to interference suppression is developed to enhance the audibility of signals corrupted by amplitude-modulated (AM) and frequency-modulated (FM) tonal interference. The suppression algorithm uses a short-time, least-squares estimation of the parameters of an AM-FM model of the time-varying tonal interference. The method, developed in a sine-wave analysis/synthesis framework, can be integrated with time and frequency modifications for further signal enhancement. Suppression is applied to single and multitone synthetic and actual AM-FM interference, the latter including man-made signals (e.g., siren interference) and those that occur naturally (e.g., biologic interference). The relative advantages and disadvantages of the sine-wave framework in contrast to a short-time Fourier transform overlap-add framework are described. The enhancement techniques are robust in a large range of environments and can be designed to preserve a random noise background. Finally, it is shown that interference suppression on multichannels prior to beamforming enhances beamformer performance.



ACKNOWLEDGMENTS

The authors acknowledge Tom Hanna and John Harvey of the Naval Submarine Medical Research Center and Stan Hollis of the Naval Undersea Systems Center for useful suggestions and feedback in evaluation of the suppression algorithms. We also acknowlegde the Lincoln Laboratory Innovative Research Program Technical Review Committee for support of this work.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS
(Continued)

LIST OF TABLES

1. INTRODUCTION

There are numerous scenarios in which a desired signal is corrupted by amplitude-modulated (AM) and frequency-modulated (FM) tonal interference. These include, for example, interfering multitonal biologics in underwater exploration, background sirens in vehicle communications, and interfering rotating machinery for machine tool diagnosis. A characteristic of the interference is that the AM and FM may be rapidly varying and thus difficult to track and remove. In addition, the interference typically may be at a higher level than the underlying signal of interest. Tracking and removing such large time-varying interference is often difficult to achieve without distorting the signal of interest.

In this report, sine-wave analysis/synthesis [1,2] is used as a framework in which to develop a new approach to interference suppression for enhancing the audibility of signals corrupted by single or multitones with AM and FM [3]. The suppression algorithm uses a short-time, least-squares estimation of the parameters of an AM-FM model of the time-varying interference. An interference signal is constructed from the estimated model parameters. Those components of the sine-wave representation of the received signal that are due to the interference are removed to form the sine-wave representation of the desired signal. Because the synthesis of the desired signal is sinusoid-based, it is straightforward also to perform signal modification such as slow-motion audio replay. This technique extends the time duration of a signal without changing its frequency characteristic and allows the listener to capture short-duration, rapidly changing events. The new approach to enhancement is being developed with the additional constraint of preserving the perceptual quality of the environment (e.g., a colored noise background) in the enhanced output to minimize the detection of falsely perceived acoustic signals. Preliminary processing of synthetic and actual interference, the latter including man-made acoustic signals (e.g., siren interference) and those that occur naturally (e.g., biologic interference), shows significant enhancement in the audibility of the desired signal.

The approach of this report differs from conventional methods of time-varying tone suppression (e.g., adaptive notch filtering [4–6]), not only in the short-time analysis/synthesis framework, but also in that these approaches were not designed with the enhanced perception of wideband acoustic signals as the objective.[1] The goal of *improved audibility* raises issues not seen when the end result is improved automatic detection or enhanced visual displays; error in parameter estimation or a measure of the degree of suppression does not illustrate the complete performance of an algorithm. For example, the residual that remains after suppression, although small, may be a perceptible artifact that can be mistaken for a signal of interest. The frequency domain framework allows control of this residual as well as flexibility in guiding the suppression algorithm in the

[1] A brief overview of state-of-the-art estimation of modulated tones and their suppression is given in Appendix A.

presence of complex backgrounds. In addition, conventional techniques lack the versatility of the approach in this report, which integrates signal modification with interference suppression.

The outline of the report is as follows: Section 2 reviews the sine-wave signal representation, demonstrates its applicability to a general class of signals, and reviews an alternate short-time overlap-add analysis/synthesis procedure. Section 3 describes the new approach to single-tone interference suppression, applies it to synthetic signals, presents an approach to background preservation, and compares sine-wave and overlap-add frameworks with respect to signal time resolution and interference suppression. Section 4 gives the extension to multitone interference, demonstrates the approach with a number of actual signals, and introduces the use of frequency guides in the suppression of complex multitone interference, including harmonic guides generated from estimates of a fundamental frequency of the interference. Section 5 describes the use of the algorithm in the context of beamforming and shows that beamformer performance improves after multichannel interference suppression. Section 6 then integrates the suppression algorithm with time-scale modification for enhancement, and Section 7 summarizes and discusses future directions.

2. SINE-WAVE REPRESENTATION OF ACOUSTIC SIGNALS

The sine-wave representation of a signal is given by a sum of sine waves with time-varying amplitudes, frequencies, and phases [1,2]:

$$s(t) = \sum_{k=1}^{N} A(t,k) \cos[\theta(t,k)] \quad , \tag{1}$$

where the amplitudes and phases for the kth sine wave are denoted by $A(t,k)$ and $\theta(t,k)$, respectively. The time-varying frequency of each sine wave is given by the derivative of the phase and is denoted by $\omega(t,k) = \dot{\theta}(t,k)$, which is sometimes referred to as the kth "frequency track." Although this model was originally formulated for speech signals, it is also capable of representing complex acoustic nonspeech signals.

2.1 Analysis/Synthesis

Using the sine-wave model of Equation (1), a discrete-time[2] analysis/synthesis system has been developed [1,2] (see Figure 1). On each analysis frame the sine-wave parameters are estimated at time samples $n = mQ$, where the frame number $m = 0, 1, 2...$, and where Q is the number of samples in the frame interval. The dependence of the sine-wave parameters on the discrete-time variable n is therefore replaced by their dependence on the frame number m, e.g., $A(n,k)$ is replaced by $A(mQ,k)$ or for simplicity by $A(m,k)$. A 3- to 10-ms frame interval has been found to produce high-quality reconstruction for most signals of interest. The analysis window (Hamming, typically 5 to 25 ms in duration) denoted by $w(n)$, is placed symmetric relative to the origin, which is defined as the center of the current analysis frame. A discrete short-time Fourier transform (STFT) is then computed over this duration with a fast Fourier transform (FFT), typically 1024 or 2048 points. The frequencies $\omega(m,k)$ are estimated by picking the peaks of the uniformly spaced (FFT) samples of the short-time Fourier transform magnitude (STFTM). The sine-wave amplitudes $A(m,k)$ and phases $\theta(m,k)$ at the center of each analysis frame are then given by the amplitude and phase of the STFT at the measured frequencies.

The first step in synthesis requires associating the frequencies $\omega(m,k)$ measured on one frame with those obtained on a successive frame. This initial step is accomplished with a nearest-neighbor matching algorithm, which incorporates a birth-death process of the component sine waves, i.e., they are allowed to come and go in time. The amplitude $A(m,k)$ and the phase $\theta(m,k)$ parameters are then interpolated across frame boundaries at the matched frequencies to upsample to the

[2]Because measurements are made using digitized sounds, sampled-data notation is used typically throughout this report.

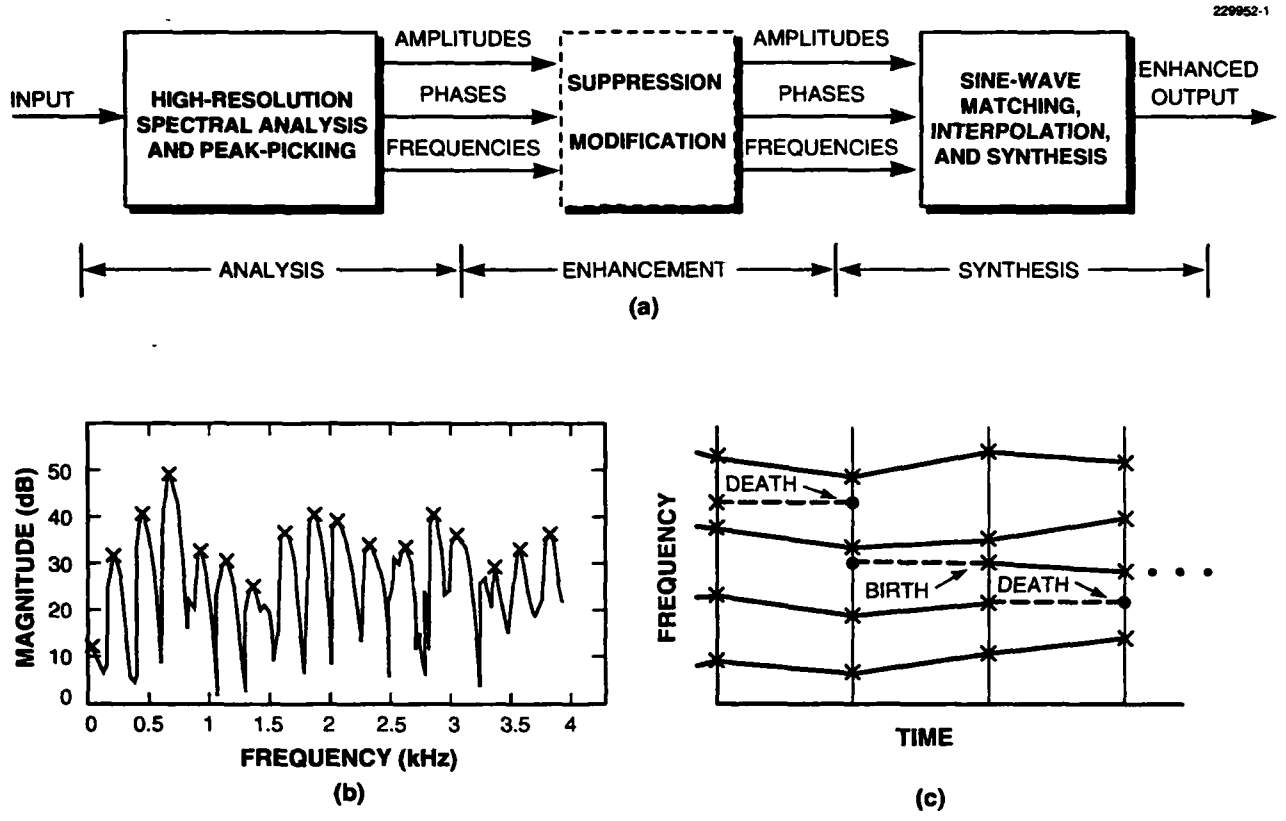


Figure 1. Sinusoidal transform analysis/synthesis system: (a) block diagram of analysis/synthesis with enhancement, (b) STFTM with sine-wave peaks, and (c) sine-wave frequency matching.

original sampling rate. The amplitude is interpolated linearly and the phase is interpolated with a cubic polynomial, the latter being done using the methods described in McAulay and Quatieri [1] and Quatieri and McAulay [2]. The interpolated amplitude and phase components are then used to form an estimate of the waveform according to Equation (1).

2.2 Application to Nonspeech Signals

The enhancement problem is concerned with two signal classes: the desired acoustic signal (i.e., the signal to be enhanced) and the unwanted background signal. Because the interest is to enhance nonspeech as well as speech sounds, about 25 signals were collected from audio recordings of complex acoustic signals (e.g., a bouncing can, a slamming book, a closing stapler). These signals were selected to have different attack characteristics and a variety of time envelopes and spectral resonances. Various synthetic and real background signals, comprising AM-FM tonal interference as well as random noise, were collected. AM-FM interference included man-made

signals (e.g., a blaring siren), biologic signals (e.g., a porpoise cry), and geologic signals (e.g., rubbing ice plates). Random background signals included white and colored synthetic noise as well as actual backgrounds (e.g., an ocean squall and an underground explosion).

Although the sine-wave analysis/synthesis is not strictly an identity, the sine-wave reconstruction of such complex acoustic signals was found to be nearly perceptually indistinguishable from the original. An example of reconstruction of an acoustic signal from a closing stapler is shown in Figure 2. To attain the time and frequency resolution required to reconstruct such signals, the duration of the analysis window $w(n)$, the number of sine-wave peaks N, and the frame interval Q are adapted to the signal type. In this example, a 7-ms analysis window, a 3-ms frame, and about 50 peaks were used. Because the window duration is typically set to obtain adequate spectral resolution, some temporal smearing can occur for short duration signals and signals with sharp attacks (as observed in Figure 2) and sometimes perceived as a mild dulling of the sound. In the reconstruction of random (background) signals, because the number of peaks may not be adequate for a noise representation, occasionally a slight (nearly imperceptible) "tonality" may be introduced.

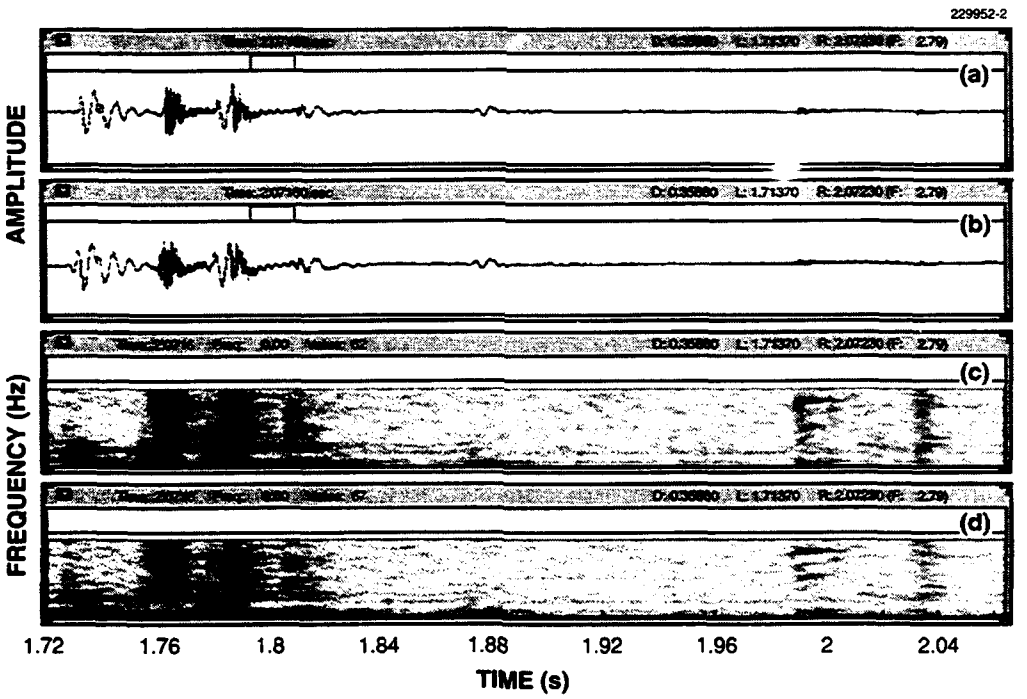


Figure 2. Sine-wave reconstruction of acoustic signal from closing stapler: (a) original, (b) reconstruction, (c) and (d) spectrograms of (a) and (b).

For a large class of signals, sine-wave analysis/synthesis is nearly a perceptual identity system, and signals are expressed in terms of a functional model describing the behavior of each of its sine-wave components. The sine-wave representation, therefore, provides an appropriate framework for developing signal enhancement techniques based on transforming each of the functional descriptors (see Figure 1).

2.3 Comparison with Overlap-Add Analysis/Synthesis

Many methods in this report can also be developed in a short-time overlap-add framework, where a discrete-time signal is represented by its STFT

$$S(mQ,\omega) = \sum_n s(n)w(n-mQ)\exp[j\omega n] \quad , \tag{2}$$

where the signal $s(n)$ is windowed with $w(n)$, the short-time analysis window, and Q is the frame interval. The sliding window and frame interval are designed for perfect reconstruction in time [7]

$$\sum_m w(n-mQ) = 1 \tag{3}$$

so that overlap-add analysis/synthesis, unlike sine-wave analysis/synthesis, is an identity. On the other hand, sine-wave analysis/synthesis gives a functional description of the underlying signal components that is not provided by the overlap-add representation.

2.4 Discussion

In reviewing sine-wave and overlap-add analysis/synthesis for signal representation, although sine-wave analysis/synthesis appears to be at a disadvantage in terms of recovering a signal, i.e., it is not (mathematically) an identity, the overlap-add method suffers from a disadvantage in its suppression capability as well as in integrability with signal modification schemes. A comparison of the overlap-add and sine-wave frameworks with respect to time and frequency resolution, as well suppression performance, is given in Section 3.5.

3. INTERFERENCE SUPPRESSION

The AM-FM tonal interference model is assumed of the form

$$i(t) = a(t)\cos[\phi(t)] \quad , \tag{4}$$

where $a(t)$ is the amplitude envelope, and where the signal frequency $\omega(t)$ is given by the derivative of the phase $\phi(t)$, i.e., $\omega(t) = \dot{\phi}(t)$. The amplitude and frequency are assumed to vary linearly over a short-time duration (e.g., 10 to 50 ms). A piecewise linear model, therefore, is assumed for the amplitude modulation

$$a(t) = A_0 + A_s t \tag{5}$$

and the phase is modeled as piecewise quadratic

$$\phi(t) = \omega_0 t + \int_{\tau=0}^{t} \omega(\tau) d\tau + \phi_0 \quad , \tag{6}$$

where ω_0 is the "carrier" frequency, $\omega(t) = \omega_s t$ with ω_s being the frequency sweep rate, and ϕ_0 is the initial phase. In practice, a single-tone interference signal does not strictly follow the piecewise linear amplitude and frequency model, but the model is sufficiently dynamic to reasonably approximate many interference signals of interest.

The received signal $r(n)$ to be processed is given in discrete time by

$$r(n) = d(n) + i(n) + b(n) \quad , \tag{7}$$

where $d(n)$ is the desired acoustic signal, henceforth referred to as the "information signal," $i(n)$ is the AM-FM interference signal [which is assumed to have a larger power level than $d(n)$], and $b(n)$ is some other background interference (e.g., white noise). Assuming for the moment that $b(n) = 0$, the sine-wave components of $d(n)$, estimated on each analysis frame as described in Section 2, can be thought of as corrupted by the STFT of $i(n)$ evaluated at the sine-wave frequencies of $d(n)$. This relation is written in complex form on the mth frame and for the kth sine wave as

$$R(m,k) = D(m,k) + I(m,\omega_k) \quad , \tag{8}$$

where $R(m,k)$ and $D(m,k)$ are the sine-wave representations of $r(n)$ and $d(n)$, respectively, and where $\omega_k = \omega(m,k)$ (with the argument m dropped for simplicity) are the sine-wave frequencies that are obtained by peak-picking the STFT magnitude of the received signal $r(n)$. It is assumed

that this one frequency set respresents the sine-wave frequencies for both $r(n)$ and $d(n)$. $R(m,k)$ can be expressed in terms of the measured sine-wave amplitude and phase

$$R(m,k) = A_r(m,k)\exp[j\theta_r(m,k)] \quad . \tag{9}$$

Likewise, $D(m,k)$ can be written in terms of the desired sine-wave amplitude and phase

$$D(m,k) = A_d(m,k)\exp[j\theta_d(m,k)] \tag{10}$$

and $I(m,\omega)$, the STFT of $i(n)$, can be expressed

$$I(m,\omega) = A_i(m,\omega)\exp[j\theta_i(m,\omega)] \quad . \tag{11}$$

For Equation (8) to strictly hold, $I(m,\omega)$ is assumed sufficiently "smooth" so as not to introduce peak frequencies that are not components of $d(n)$. This smoothness constraint is illustrated in Figure 3, where an acoustic signal from a bouncing can has been added to an FM chirp signal; the Fourier transform magnitude of a short segment (25 ms) shows that the main and sidelobes of the interference are smooth relative to the spectral magnitude of the can, which is characterized by a rapidly varying spectrum.

Interference suppression in the context of sine-wave analysis/synthesis is accomplished by removing the interference contribution in the sine-wave representation of the received signal (see Figure 4). Two different approaches are considered: *magnitude-only suppression*, which removes the interference contribution to the sine-wave amplitude and leaves the phase of $R(m,k)$ intact; and *complex suppression*, which removes the sine-wave amplitude and the phase contributions due to the interference.

3.1 Magnitude-Only Suppression

In magnitude-only suppression, the estimate of the sine-wave amplitudes and phases of the desired signal is given by [3]

$$\hat{A}_d(m,k) = A_r(m,k) - \hat{A}_i(m,\omega_k) \tag{12}$$

$$\hat{\theta}_d(m,k) = \theta_r(m,k) \quad , \tag{13}$$

---

[3]Because $\hat{A}_i(m,\omega_k)$ is an estimate, it is possible that $\hat{A}_d(m,k)$ may be negative; and because negative sine-wave amplitudes are not meaningful, these values of $\hat{A}_d(m,k)$ are set to zero.

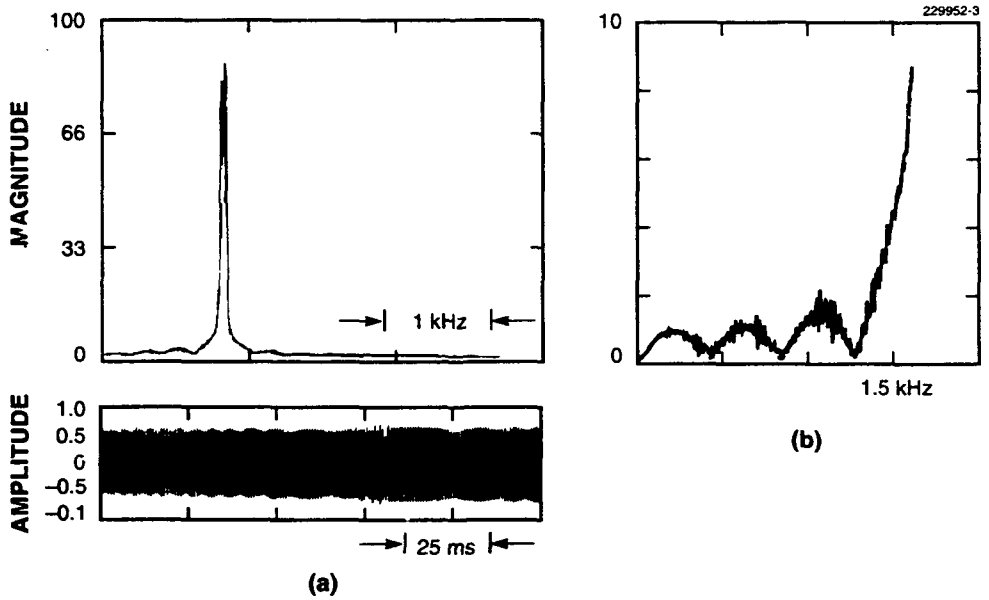


Figure 3. Smoothness constraint on spectral interference with respect to information signal: (a) waveform and spectrum and (b) blowup of spectrum. (Information signal is a bouncing can.)

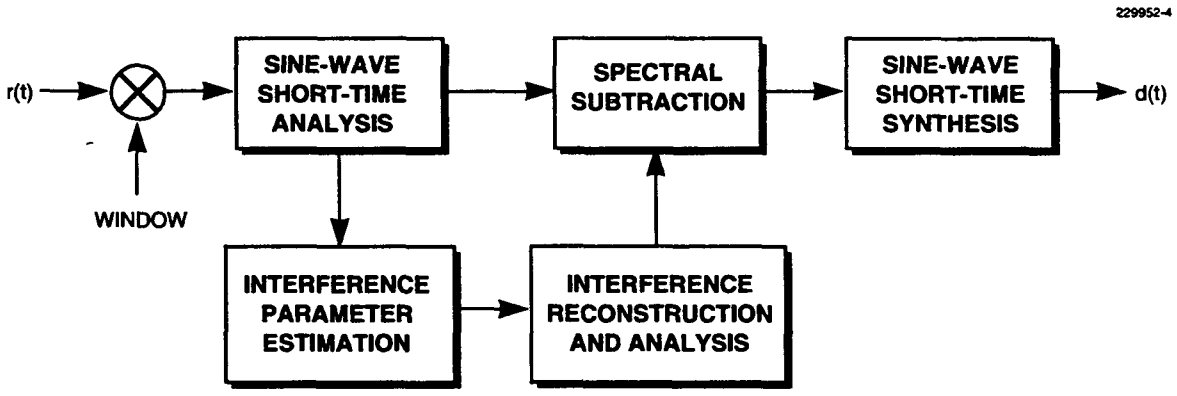


Figure 4. Approach to suppression of AM-FM tonal interference.

where "hat" denotes estimate. With the amplitude and phase estimates in Equations (12) and (13), an estimate of the desired signal can be made using the sine-wave synthesis of Section 2.

Interference suppression requires that an estimate of the sine-wave magnitude contribution from the interference signal, $\hat{A}_i(m,\omega_k)$, be computed for each frame. From Equations (4), (5), and (6), the interference signal on frame m is modeled in discrete time as

$$i(n) = A_o \cos(\omega_o n + \omega_s \frac{n^2}{2} + \phi_o) \tag{14}$$

with $n = 0$ corresponding to the center of the analysis window [where reference to frame m in (Equation (14) is implicit], where the sampling period is equal to unity, and where $a(t)$ in Equation (5) has been made piecewise constant over each frame. There are four unknown parameters of $i(n)$: the amplitude A_o, the carrier frequency ω_o, the initial phase ϕ_o, and the sweep frequency ω_s. Because the AM-FM signal is assumed to change linearly and "slowly" over the duration of the symmetric analysis window, and because the interference is assumed to dominate the desired signal, estimates of A_o, ϕ_o, and ω_o are obtained at the maximum in the magnitude of the STFT of the received signal, $|R(m,\omega)|$ [8]. The sweep frequency ω_s is estimated by tracking ω_0 over successive frames; specifically, the estimate $\hat{\omega}_s$ is the slope of a line that is a least-squares fit to the successive values of ω_o. The STFTM of the AM-FM interference estimate is then evaluated at the measured sine-wave frequencies ω_k to form the estimated sine-wave amplitudes due to the interference $\hat{A}_i(m,\omega_k)$. These amplitudes are subtracted from the received signal to form the sine-wave representation of the information signal.

3.2 Complex Suppression

An implicit assumption in the magnitude-only suppression algorithm is that the measured peak amplitudes are the sum of the peak amplitudes of the information signal and samples of $|I(m,\omega)|$; this assumption is an approximation due to the complex nature of the Fourier transform. This approximation and the use of the phase of the received signal in the reconstruction introduces distortion in the estimated information signal; ideally, then, a complex subtraction should be performed.

In complex suppression the sine-wave amplitudes and phases are obtained by a vector subtraction

$$\hat{D}(m,k) = R(m,k) - \hat{I}(m,\omega_k) \quad , \tag{15}$$

where hat denotes the estimates of the respective quantities in Equation (8). As in magnitude-only suppression, an estimate of the parameters of $I(m,\omega)$ can be obtained via the maximum of $|R(m,\omega)|$. The complex nature of the subtraction, however, prohibits an accurate suppression with these coarse estimates (see Appendix B).

To account for this sensitivity, the error function defined by

$$\epsilon = \sum_n [w(n)[r(n) - i(n)]]^2 \quad , \tag{16}$$

where $w(n)$ is the analysis window, is minimized over the parameters of the model for $i(n)$ given by

$$i(n) = (A_0 + A_s n)\cos[\omega_0 n + \omega_s \frac{n^2}{2} + \phi_0] \quad , \tag{17}$$

where a linear sweep is incorporated back into the amplitude envelope to improve the accuracy of the model. This error minimization approach was selected for parameter estimation because similar estimation methods are known to give good performance for the linear FM/constant amplitude case [8]. The highly nonlinear problem of minimizing ϵ with five free parameters can be solved with various well-established iterative methods. The Powell method was chosen for its computational ease and relatively rapid convergence [9,10]. The starting point in the Powell iterative method uses the coarse parameter estimates derived from magnitude-only suppression. The iteration ends when the change in the mean-squared error falls below a fixed threshold; for the signals of interest, typically 5 iterations (with a maximum of about 20) are required in the Powell algorithm. (See Appendix C for further discussion of this approach and alternate methods that have been considered for least-squares estimation.)

Although the least-squares error approach has been motivated by complex suppression, it can also be used in refining the magnitude-only subtraction technique. Specifically, the coarsely estimated parameters used in Section 3.1 can be replaced by the (iteratively) refined estimates. This approach to magnitude-only supression is considered further.

3.3 Examples

Figure 5 shows the result of the complex suppression algorithm applied to multiple bounces of a bouncing can with an AM-FM interference at about a 25-dB interference-to-signal ratio (ISR);[4] the can is barely audible in the presence of the interference. The interference signal used in this experiment is a tone with a sinusoidally varying instantaneous frequency $\omega(t) = 1500 +$

[4]ISR is defined by measuring the average power in the information signal over its duration and dividing this result into the average power of the interference. Defining the "duration" of a transient signal is difficult, as for example, a closing stapler or a bouncing can. Thus the signal averaging was performed only when the instantaneous power (measured using a sliding window of length of 1 ms) of the signal exceeded a threshold of 10% of the maximum instantaneous power.

$400\sin[2\pi(0.532)t]$ comprising a center frequency of 1500 Hz with a swing of 400 Hz and a maximum slope of about 1500 Hz/s; and a sinusoidally varying amplitude $A(t) = 1 + 0.2\sin[2\pi(0.617)t]$ comprising a constant of unity with a swing of 0.2 and a maximum amplitude slope of about 0.6/s. Because these modulations were selected to avoid regularities in the waveform, and because this interference signal does not strictly follow the short-time linear assumption, it provides a good test of the suppression algorithm. A 10-ms Hamming window, a 4-ms frame, and a 2048-point DFT were used; these parameter values are used throughout this report unless otherwise indicated.[5] Suppression is performed with little change in the quality of the falling can with a resulting slight "whishing" residual from the interference. The suppression ratio (defined as the average power in the interference before suppression divided by the average power in the interference residual after suppression) for this case is about 40 dB so that the resulting interference residual is below the transient by about 15dB. The signal of interest, almost imperceptible in the original signal, is clearly audible in the processed version.

A closer view of the fine time structure of the process is shown in Figure 6, which compares applying magnitude-only and complex suppression with a different test signal consisting of the AM-FM interference added to an information signal generated from a closing stapler about 25 dB below the interference. This example illustrates that complex suppression can provide a more accurate reconstruction of the information signal, but as will be shown, at the expense of a larger interference residual.

Another example (Figure 7) demonstrates the robustness of the complex suppression algorithm in a complicated background. In this example the interfering signal is a synthetic linear-FM chirp with a frequency sweep of 1000 Hz/s, and the signal of interest is an acoustic signal from a bouncing wrench. The background consists of an ocean squall, a multitonal whale cry, and ocean noise. In removing the interfering chirp, the falling wrench is enhanced while the complex background has been preserved, barring the spectral nulls at chirp center frequencies. This spectral nulling effect, the robustness of the algorithm, as well a comparison of the magnitude-only and complex suppression methods are addressed more quantitatively in Sections 3.4 and 3.5.

3.4 Performance

3.4.1 Suppression and Signal Clarity

Both magnitude-only and complex suppression provide a substantial reduction of the interference signal with the perceptual character of the information signal essentially preserved. A

[5]These parameters were empirically selected to trade-off suppression residual for reconstruction fidelity of a class of information signals with fast attacks and short duration such as a bouncing can. Further discussion of the selection of the analysis window is given in Appendix D.

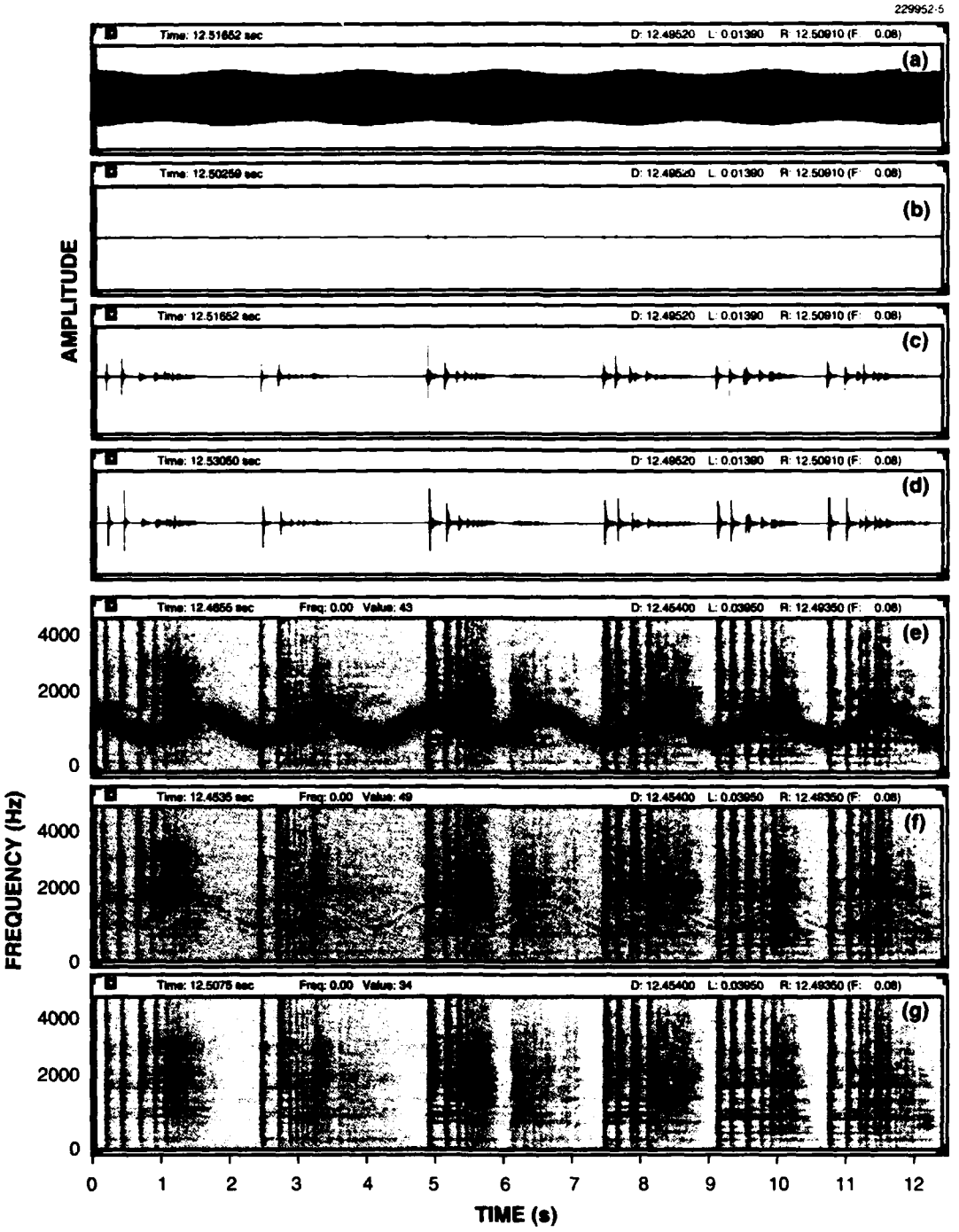


Figure 5. Recovery of weak signal dominated by AM-FM tonal interference: (a) original signal plus interference; (b) recovered signal; (c) recovered signal magnified × 15; (d) original information signal; (e), (f), and (g) spectrograms of (a), (b), and (d), respectively.

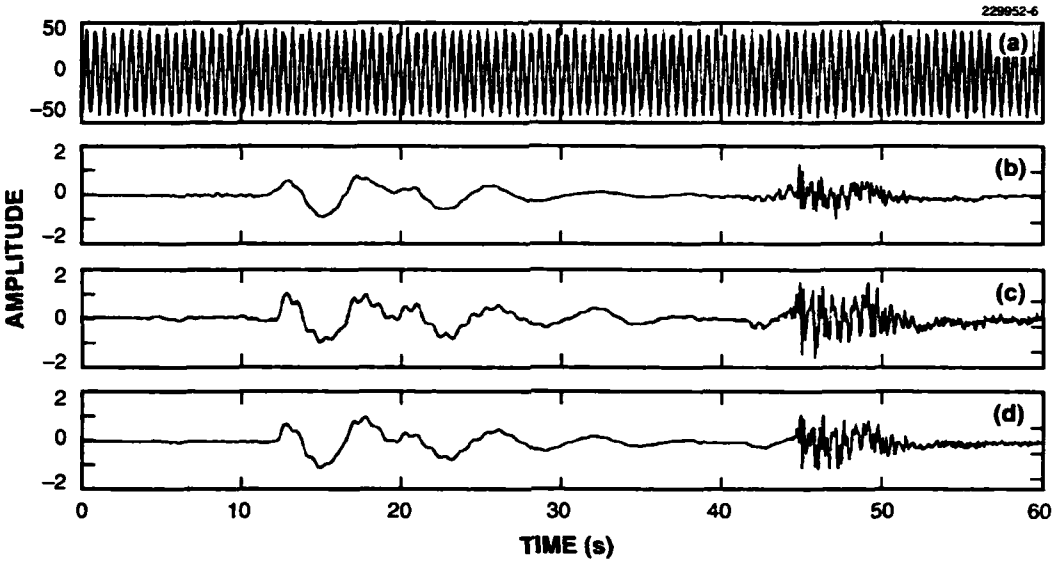


Figure 6. Example of interference suppression: (a) information signal obscured by AM-FM interference, (b) result of magnitude-only suppression; (c) result of complex suppression, and (d) original information signal.

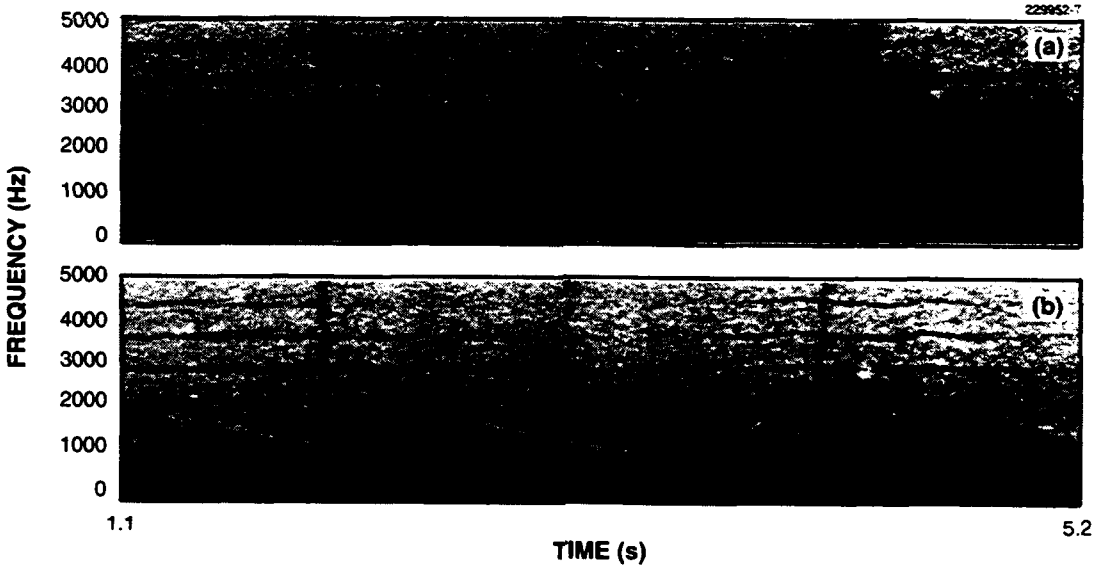


Figure 7. Preservation of complicated background: (a) spectrogram of interfering chirp with whale, squall, and wrench signal and (b) processed (a).

quantitative measure of suppression is the *suppression ratio*, earlier defined as the ratio of the interference power before suppression to the interference power remaining after suppression. Table 1 gives the suppression ratios measured by performing suppression on an interference signal with no information signal present. For completeness, the suppression measurements were made for magnitude-only and complex subtraction with both *coarse* (i.e. from the maxima of $|R(m,\omega)|$) and *refined* parameter estimates [i.e., from error minimization using Equations (16) and (17)]. The interference signal used in this experiment is the preceding tone with a sinusoidally varying instantaneous frequency $\omega(t) = 1500 + 400\sin[2\pi(0.532)t]$ and a sinusoidally varying amplitude $A(t) = 1 + 0.2\sin[2\pi(0.617)t]$. Table 1 shows that the magnitude-only method provides greater suppression than the complex method. This is not surprising because the former clips negative spectral regions in obtaining the estimate $\hat{A}_d(m,k)$ in Equation (12). The refined estimation scheme improves on the coarse estimation for both suppression methods.

TABLE 1

Suppression Performance

Interference Parameter Estimation Method	Coarse		Refined	
Suppression Method	**Magnitude**	**Complex**	**Magnitude**	**Complex**
Suppression ratio	31.6 dB	26.0 dB	51.6 dB	38.5 dB
Subjective suppression	3rd	4th	1st	2nd
Information signal clarity	4th	2nd	3rd	1st

The results of an informal listening test are also listed in Table 1, based on the judgment of interference reduction and clarity of the information signal after suppression. One test signal was created by adding the acoustic signal from a bouncing can to the interference of the previous experiment; in a second test signal the interference was added to the response of a closing stapler. In both cases the interference signal power level was about 25 dB higher than the information signal, which is virtually inaudible. Two listeners were asked to rate the interference suppression and the clarity of the estimated information signal on a scale of 1 to 4. Ratings were averaged over listeners and test signals and then rank ordered. The perceived reduction in interference follows the measured suppression ratios. The table also shows that the clarity of the information signal is better maintained by complex suppression, regardless of the method used to estimate the parameters of the interference. Because interest is primarily in enhancing the detection and discrimination of the information signal, its clarity after suppression may be more important to the listener than the amount of suppression. On the other hand, the residual interference, which is perceived as a

background modulated whishing, may be misinterpreted as an information signal. Determining the optimal trade-off between suppression and clarity requires more extensive evaluation.

3.4.2 Robustness

As a demonstration of the robustness of the algorithm, the suppression obtained from the least-mean-squared estimation (for complex suppression) in the presence of background noise [i.e., $b(n)$ in Equation (7)] is shown in Figure 8. The interference signal used in this experiment is a tone with a sinusoidally varying frequency $\omega(t) = 1500 + 400\sin[2\pi(0.532)t]$ and a constant amplitude.[6] The parameter estimation technique was found to be robust at low interference-to-noise ratios (INRs). The suppression ratio was determined by measuring the interference parameters in the presence of noise, suppressing the original interference signal (without noise present), and then comparing the power in the interference signal before and after suppression. Although the suppression ratio drops as the INR increases, the perceived interference residual in noise is removed at low INRs. In addition to measurements of complex suppression, Figure 8 also shows that magnitude-only suppression (using refined estimation) is greater than that from complex suppression; magnitude-only suppression also has an advantage with respect to computation because iterations (for coarse estimation) and phase computations are not required. These advantages are obtained at the expense of larger distortion of the information signal.

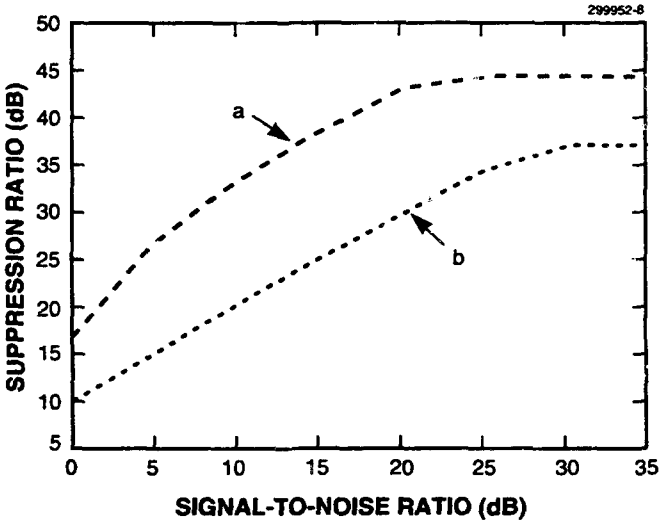


Figure 8. Suppression ratio as a function of SNR: (a) magnitude-only and (b) complex.

[6] A constant amplitude was selected because in this test a constant signal-to-noise ratio (SNR) is desirable; with AM, SNR changes with time.

3.5 Preservation of Background

In performing interference suppression in the presence of background noise $b(n)$, it is important that the perceived character of the background be maintained to minimize false detection of information signals. Magnitude-only and complex suppression largely preserve background, both synthetic (e.g., white noise) and real (e.g., an ocean squall). In the region of the time-varying AM-FM tone, however, a spectral null is formed. The problem is that the least-squares parameter estimation represented by Equations (16) and (17) yields a biased estimate of the background spectrum [11], forcing it to zero in the vicinity of the interfering chirp frequency; the suppression algorithm thus nulls the spectrum of the received signal in this region (as seen in Figure 7).[7] As the duration of the analysis window decreases, the region over which the spectrum is nulled increases; this nulling may be exacerbated by the the accuracy of the interference parameter estimates decreasing as the window length decreases (see Appendix D). This notch follows the instantaneous frequency of the interference; therefore if the background is broadband, the notch is perceived as an FM modulation that can be mistaken as an information signal. Figure 9 illustrates an example of an FM notch placed in the spectrogram of the example in Figure 5 when white noise is added to the background. This section presents two heuristic approaches to reduce the spectral notch without degrading suppression performance.

Because the spectrum of the interference at its peak frequency tends to swamp smaller peaks in its vicinity, few sine-wave peaks are picked in this region. One approach to recovering sine waves, and perhaps reducing unwanted modulation introduced by the spectral null, is to reconstruct a signal using a set of new sine waves, different from those measured on the received signal and obtained after applying suppression. Using complex suppression, this recovery can be accomplished by finding spectral peaks in the STFTM given by $|R(m,\omega) - \hat{I}(m,\omega)|$ and then determining the sine-wave representation of the information signal plus background, $d(n) + b(n)$. Although the method reduces the time-varying spectral notch, an FM whishing is nevertheless heard in the residual and correlates with a visible (although somewhat reduced) notch in the spectrogram. The original spectral bias is not fully removed by this approach.

A second approach applies a spectral compensation to the data based on the assumption of a slowly varying background. Assuming for the moment that the information signal is not present, an estimate of the background spectrum on the mth frame is obtained by averaging the squared STFTM, i.e., averaging the periodogram. The spectral density of the background is estimated as

$$\hat{B}(m,\omega) = \alpha\hat{B}(m-1,\omega) + (1-\alpha)|R(m,\omega)|^2 \quad , \tag{18}$$

---

[7]The least-squares method of suppression, which under certain conditions (e.g., a Gaussian noise assumption and constant amplitude and frequency) is equivalent to maximum likelihood spectral estimation, is biased. Specifically, the estimator finds the maximum value in the spectrum [11] that when subtracted yields a zero residual at the maximum location; hence a notch.

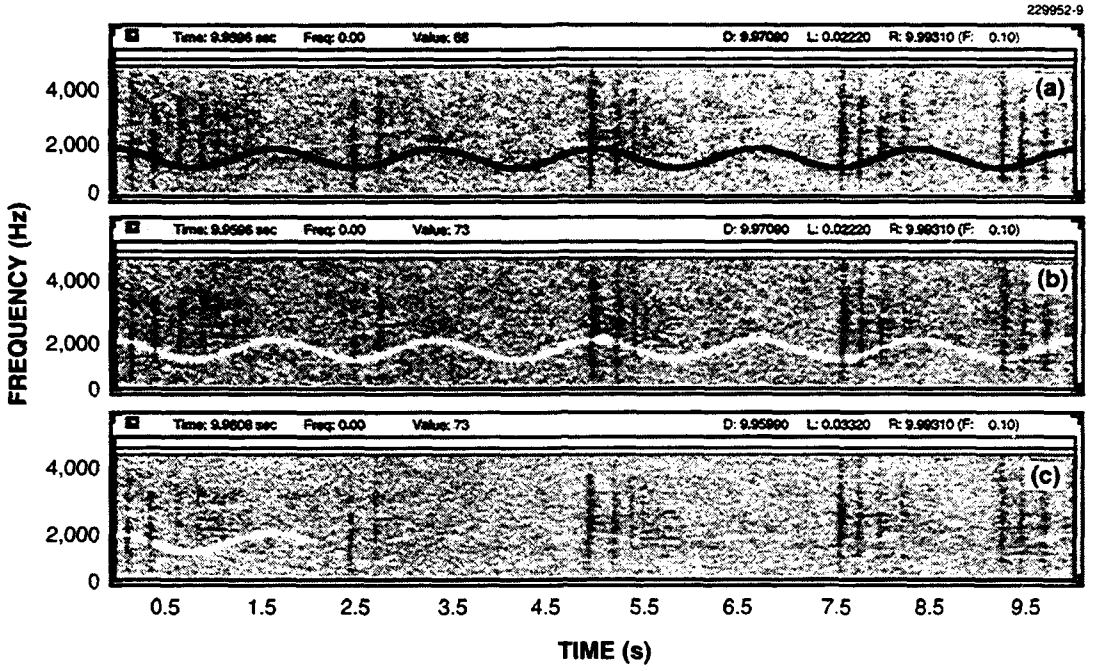


Figure 9. AM-FM tone interference with bouncing can in noise: (a) original, (b) processed, and (c) processed with spectral compensation.

where $R(m,\omega)$ is the STFT of the received signal $r(n)$ and α is a smoothing constant. This method is similar to the Welch method of spectral estimation [12]. When the background is a stationary random process, it can be shown that the expected value of $B(\omega; mL)$ is a smooth version of the desired spectrum

$$E[B(\omega; mL)] = \gamma \int_0^{2\pi} B_i(\tau)|W(\omega - \tau)|^2 d\tau \quad , \tag{19}$$

where $B_i(\omega)$ is the underlying spectral density of the background, $W(\omega)$ is the Fourier transform of the window $w(n)$, and γ is a function of both the window length and the smoothing constant α. When the interference is present and a spectral notch arises from suppression, the spectral density of the notched spectrum can be estimated as

$$\hat{N}(m,\omega) = \beta \hat{N}(m-1,\omega) + (1-\beta)|R(m,\omega) - \hat{I}(m,\omega)|^2 \tag{20}$$

under the assumption that the spectral notch is slowy varying. A compensation filter is then formed as

$$\begin{aligned} C(m,\omega) &= \hat{B}(m,\omega)^{1/2} - \hat{N}(m,\omega)^{1/2} \; for \; |\omega - \hat{\omega}_o| \leq \delta \\ &= 0 \; for \; |\omega - \hat{\omega}_o| > \delta \quad , \end{aligned} \tag{21}$$

where δ defines the region over which the compensation is applied, and where $\hat{\omega}_o$ is an estimate of the chirp center frequency. The filter $C(m,\omega)$ is characterized by a single-peak spectrum at $\hat{\omega}_o$, which is considered the complement of the notch. Compensation then forms a modified spectral magnitude

$$|\tilde{D}(m,\omega)| = |\hat{D}(m,\omega)| + C(m,\omega) \quad , \tag{22}$$

which has the effect of "filling in" the spectral hole due to suppression. The phase (which is dominated by the smooth phase of the chirp in the vicinity of the notch) is left intact by this operation.

The periodogram averaging results in a smooth estimate of the background density; and because the resulting phase is smooth, so is the phase of the compensated STFT in the neighborhood of the notch. A smooth phase, however, is not consistent with a typical random background. One approach to ensure a noise-like characteristic of the modified complex spectrum $\tilde{D}(m,\omega)$ is to impart phase randomization in the frequency region $|\omega - \hat{\omega}_o| \leq \delta$.[8] The phase of the resulting STFT is given by

$$\begin{aligned} \angle\tilde{D}(m,\omega) &= \pi\epsilon \; for \; |\omega - \hat{\omega}_o| \leq \delta \\ &= \angle\hat{D}(m,\omega) \; for \; |\omega - \hat{\omega}_o| > \delta \quad , \end{aligned} \tag{23}$$

where ϵ is a random number falling uniformly in the interval $[-1,1]$.

An example of the removal of spectral bias is shown in Figure 10 for a steady tone in noise, where the tone onset occurred two seconds into the noise. In this example the background spectral estimator was applied only up to the onset of the tonal interference, at which point the compensation filter was activated. The average of the spectral slices illustrates that the resulting spectrum in the neighborhood of the notch is consistent with the surrounding background spectrum. Figure 11 illustrates another example of compensation with the AM-FM tone in noise that was illustrated

[8] An alternate approach synthesizes a signal by passing white noise through a linear system with transfer function $|C(m,\omega)|$ and then adding the resulting waveform to the signal derived from the suppression algorithm.

in Figure 9. As before, the background estimation occurs prior to the tone at which point the compensation is activated. The spectrogram shows that the new procedure effectively eliminates the spectral hole; furthermore, the perception of an FM residual is not present after compensation. Finally, in Figure 9(c) the method is applied to the example of Figure 9(a),(b) of a notch occurring in the spectrogram of the enhanced can in noise. When applying the compensation algorithm to this signal, the background spectral estimate was run in the time interval 1.5 to 2 s, which is a region roughly free of bounces of the falling can. The compensation filter was activated at 2 s, at which point the perceived FM effectively vanishes. Nevertheless, a small residual (visible) notch remains because the background estimator is somewhat influenced by the presence of the information signal; ideally only background regions should be used for updating the background spectral estimate. In addition, the presence of the information signal influences the spectral estimate of the notch. In this case, the AM-FM tonal interference and information signal occasionally coincide in frequency; hence, the background spectral extrapolation (in time) used to compensate for the notch should extrapolate (in frequency) from the spectrum of the information signal and not from the noise background. Such adaptive schemes require detection of the information signal and background regions.

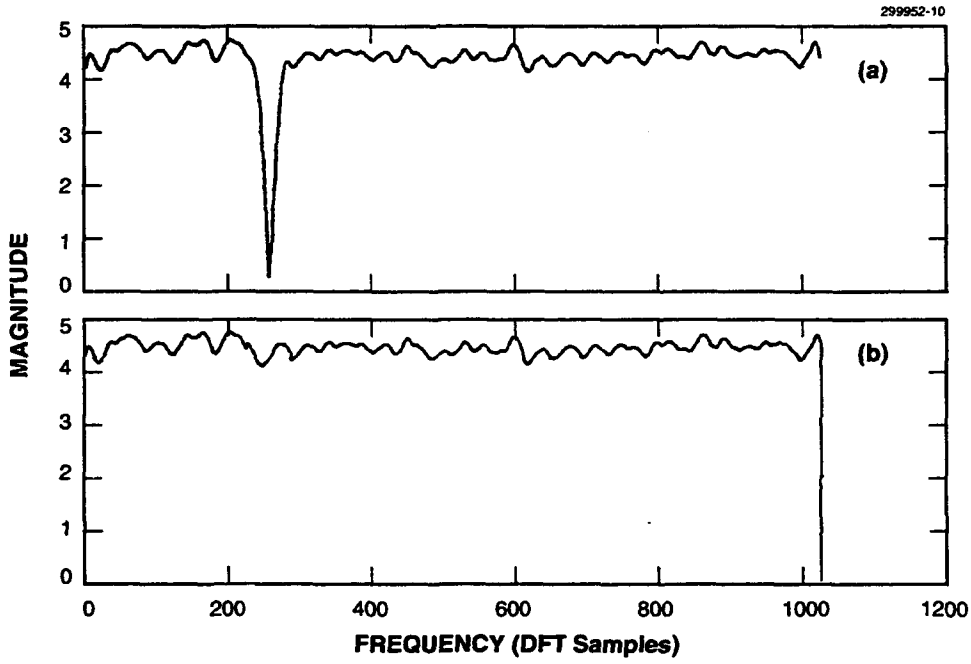


Figure 10. Removal of spectral bias for steady tone in noise: (a) average spectrum after suppression and (b) average of (a) with compensation.

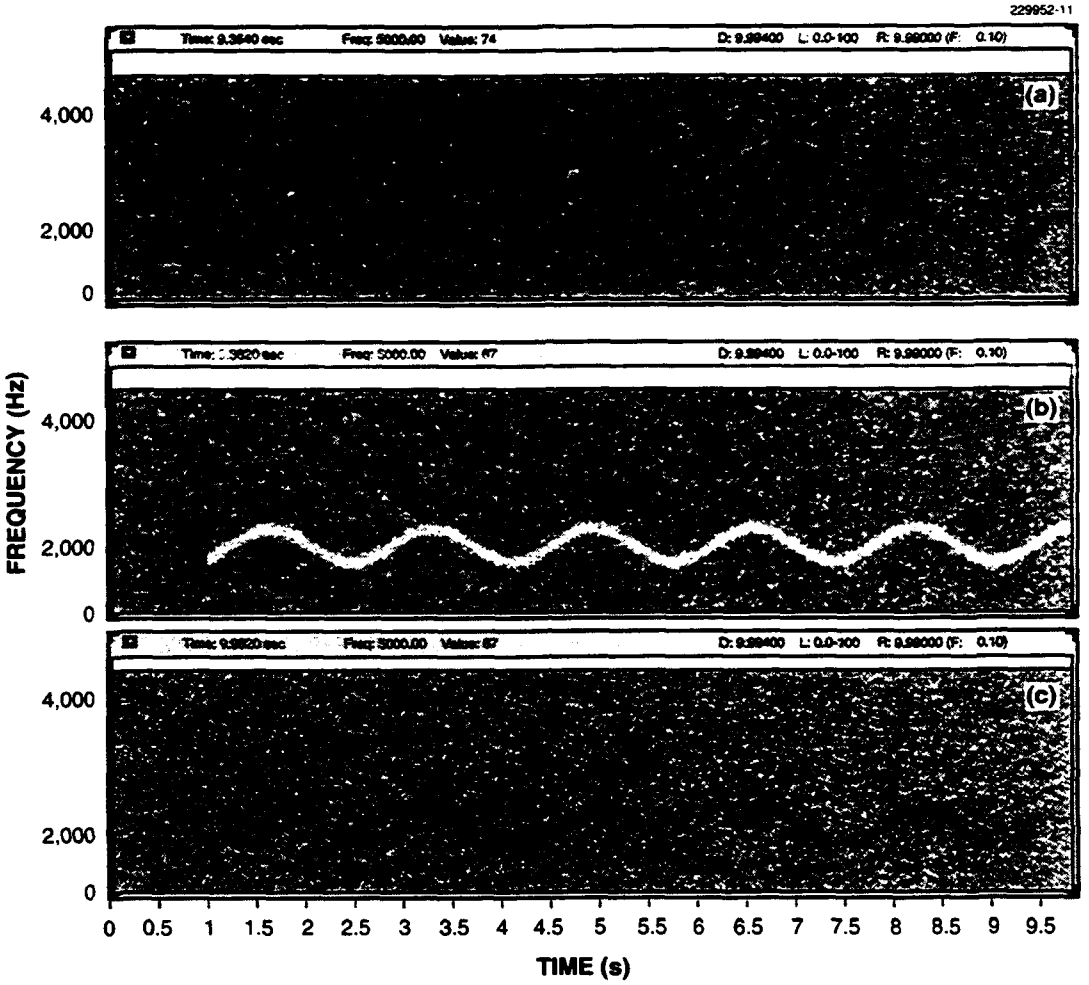


Figure 11. Removal of spectral bias for AM-FM tone in noise: (a) original, (b) processed, and (c) processed with compensation.

3.6 Comparison with Overlap-Add

In the overlap-add framework, suppression in Equations (12), (13), and (15) is performed over the full spectral bandwidth rather than over only the sine-wave frequency components. In complex suppression in particular, the modified STFT is given by

$$\hat{D}(m,\omega) = R(m,\omega) - \hat{I}(m,\omega) \quad , \tag{24}$$

where hat denotes estimates of the respective quantities. The modified short-time transform is then inverted and the resulting segments overlapped and added to form the enhanced signal. Under the perfect reconstruction condition Equation (3), the information signal would be exactly recovered when the interference is completely removed. The overlap-add framework, therefore, is expected to yield greater time resolution in the information signal, as well as greater background fidelity[9] than the sine-wave framework. Although this advantage generally holds, the overlap-add method achieves these gains at the expense of less interference suppression. The two frameworks are compared with respect to the degree of suppression using the suppression ratio, as well as with respect to the fidelity of the estimated information signal using a segmental SNR.

3.6.1 Degree of Suppression

A series of experiments was performed to compare the degree of suppression of the two algorithms. Figure 12 shows the suppression ratio from the sine-wave suppression (SWS) and overlap-add suppression (OLAS) algorithms as a function of window length.[10] Three different interference signals are considered in increasing order of complexity: a steady continuous wave (CW) tone, a linear AM-FM chirp, and a sinusoidally varying AM-FM tonal interference. Figure 12 illustrates that the sine-wave framework generally provides a greater degree of suppression than its overlap-add counterpart. For both the CW tone and linear AM-FM chirp the suppression ratio increases with window length; in these cases the accuracy of the parameter estimates of the linear AM-FM model under study increases with window length. For long window durations, the sinusoidally varying AM-FM chirp, however, violates the linear assumptions, and the accuracy of the parameter estimates decreases. The selection of window length, then, is a function of the data type; additionally, in the case of the sine-wave framework, the fidelity of the recovered information signal must be considered as well because longer windows can reduce time resolution.

One plausible explanation for the improved suppression within the sine-wave framework is based on the viewpoint that the interference estimate is most accurate at the center of the analysis window. This idea is consistent with the sine-wave analysis/synthesis strategy that estimates sine-wave parameters at the window center. Overlap-add synthesis, on the other hand, uses the entire analysis window for the reconstruction, permitting a poor interference estimate at the window edges (e.g., the Hamming window trails off). In addition, within the sine-wave framework spectral subtraction is performed only at the spectral peaks, where phase estimates are most reliable.

---

[9]Recall that sine-wave reconstruction can impart a slight tonality in the background.

[10]In the presence of noise, the suppression ratio is not the only criterion in selecting a window length. A second consideration is the power removed, because this value partially reflects background distortion. A study of these trade-offs is given in Appendix D for suppression performed within the overlap-add framework, although the issues are similar for the sine-wave framework.

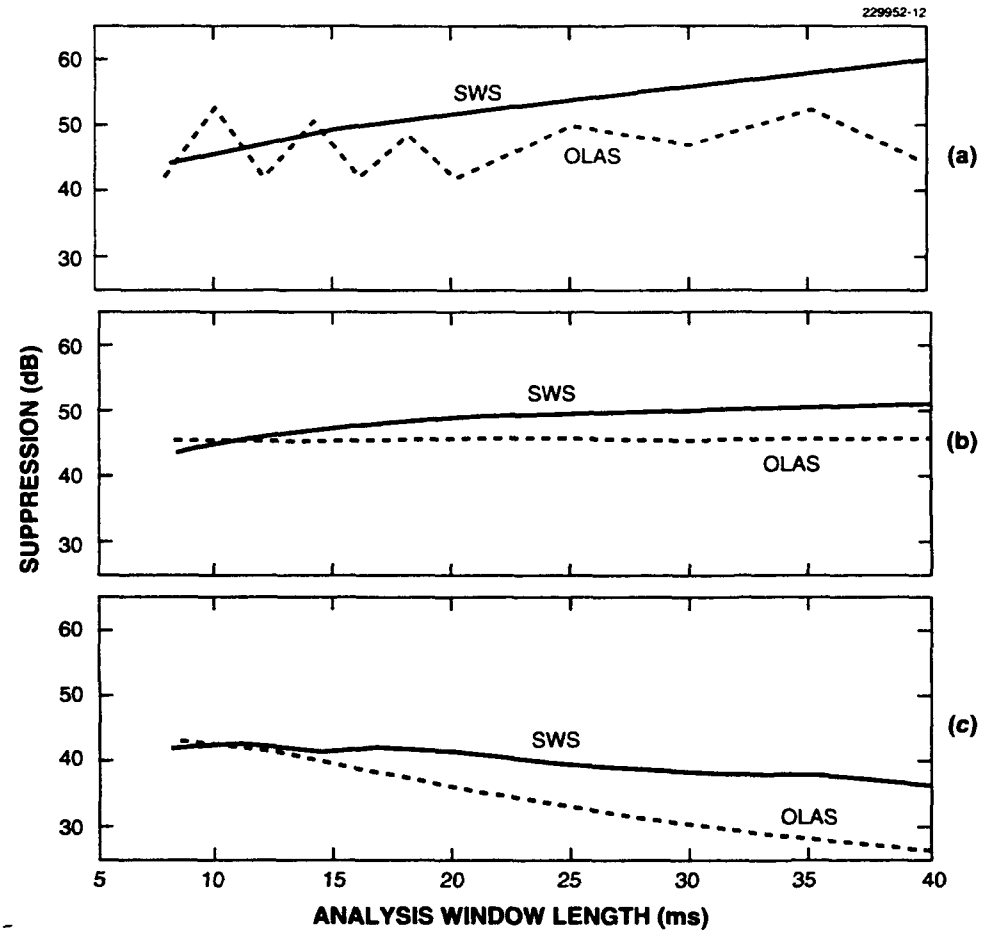


Figure 12. Comparison of SWS and OLAS for (a) CW tone, (b) linear AM-FM chirp, and (c) sinusoidal AM-FM chirp.

Finally, sine-wave synthesis provides signal continuity over consecutive frames through phase and amplitude interpolation while overlap-add may suffer from discontinuities at frame boundaries.

3.6.2 Information Signal Fidelity

To evaluate the capability of the algorithms to preserve the information signal, *segmental* SNR was used as a measure of signal distortion. Segmental SNR is the signal energy of the information signal divided by the mean squared difference between the original information signal (averaged over many segments) and its estimate after suppression. The interference signal in this study is the sinusoidal AM-FM tone used earlier. Table 2 shows that the overlap-add framework provides greater segmental SNR for three different information signals. Included for reference is the segmental SNR from sine-wave analysis/synthesis without interference and hence without suppression, providing an upper bound on the accuracy of the sine-wave reconstruction. For the stapler and the wire-wrap tool, overlap-add synthesis has a higher segmental SNR than the upper bound for the sine-wave system. The segmental SNR shows that the overlap-add scheme yields greater fidelity of the information signal than sine-wave analysis/synthesis, correlating with the result of informal listening tests.[11] Of course, this result must be tempered by the lower degree of suppression when using the overlap-add framework.

TABLE 2

Information Signal Reconstruction

Signal	Sine-Wave Synthesis (No Interference)	Sine-Wave Synthesis (Magnitude-Only Subtraction)	Sine-Wave Synthesis (Complex Subtraction)	Overlap-Add Synthesis (Complex Subtraction)
Bouncing can	11.60	3.10	2.80	13.60
Stapler	8.70	2.70	1.00	8.10
Wire-wrap tool	9.30	3.70	3.40	12.70

[11]A problem with the segmental SNR measure is that it reflects the interference residual as well as signal distortion. Although in these examples signal distortion appears to dominate over interference residual, the accuracy of the comparison must, in general, be considered with care. For example, comparison of sine-wave magnitude-only and complex suppression in Table 2 does not reflect the difference in signal clarity.

3.7 Discussion

An exhaustive comparative study of different approaches to suppression involves a complex perceptual space: the fidelity of the information signal (e.g., duller attacks), the extent and nature of the interference residual (e.g., an FM notch), and the fidelity of the background (e.g., tonality). Selecting a metric to account for all three remains an open question. The ultimate judge is the human listener's ability to detect and discriminate signals after suppression; more formal listening tests are necessary to make a complete evaluation.

4. MULTITONE INTERFERENCE SUPPRESSION

Interference signals of interest are typically characterized by multiple AM-FM tones; for example, nonlinear distortion in the sound generation process may create harmonics of a fundamental frequency or introduce new frequencies. The multitone interference $i(n)$ is modeled by

$$i(t) = \sum_{m=0}^{M} a_m(t)\cos[\varphi_m(t)] \quad , \tag{25}$$

where M is the number of tones, and $a_m(t)$ and $\phi_m(t)$ are each represented, respectively, by the linear and quadratic functions of Equations (5) and (6).

As with single-tone suppression, either a magnitude-only or complex suppression can be performed. In magnitude-only suppression, a generalization of the single-tone case entails estimating the STFTM of the interference using the spectral peaks (the highest M peaks), and then forming a spectral magnitude subtraction as a generalization of Equation (12). Successively performing the subtraction in order of increasing magnitude is efficient, but a problem with this approach is that the resulting spectral amplitude may be truncated to zero in multiple spectral locations whenever the difference in Equaton (12) becomes negative. Consequently, in losing a large portion of its spectral energy the information signal can be severely distorted.[12] For this reason complex suppression is selected for multitone interference.

In this section the complex suppression method of Section 3.2 is extended to the multitone problem. Results similar to those for the single tone are obtained with a variety of multitone interference signals, both in interference suppression and information signal clarity.

4.1 Complex Suppression

In complex multitone suppression, the error function defined by

$$\epsilon_k = \sum_n [w(n)[r(n) - i(n)]]^2 \quad , \tag{26}$$

where $w(n)$ is the analysis window, is minimized over the parameters of the model for $i(n)$ given by Equation (25). This highly nonlinear, multivariable problem is simplified by minimizing Equation (26) with respect to the parameters of one component of (25); i.e., the new error function becomes

[12] An alternate approach simultaneously estimates all tones rather than delete them iteratively, which may reduce distortion because the spectral magnitude is constructed once and occurs after multitone addition.

$$\epsilon = \sum_n (w(n)[r(n) - a_k(n)\cos[\phi_k(n)]])^2 \quad , \tag{27}$$

which when minimized yields a solution of the form

$$\hat{i}_k(n) = \hat{a}_k(n)\cos[\hat{\phi}_k(n)] \quad . \tag{28}$$

The estimate of the kth interference tone $\hat{i}_k(n)$ is then subtracted from $r(n)$ to form

$$\hat{r}_k(n) = r(n) - \hat{i}_k(n) \quad , \tag{29}$$

and the minimization is repeated with $\hat{r}_k(n)$ as the received signal. As with single-tone suppression the iterative Powell method of minimization is used (see Appendix C for further discussion).

Without constraints, the global mininum of Equation (27) may not occur in the least-squares minimization and thus may not yield the largest tonal component (or perhaps not any tonal component) of $i(n)$; in general, this procedure may be stymied by local minima. A means to avoid unwanted minima is to initialize the minimization procedure by a guess near the largest interference component. When the interference signals are quasi-harmonic in nature or with predictable frequency relations (as from nonlinear distortion), an estimate of the fundamental frequency (or few primary frequencies) helps guide the frequency search of tones belonging to the interference, thus reducing the possibility of achieving undesired local minima. Some of these signal scenarios, as well as the performance of the suppression algorithm, are illustrated next.

4.2 Performance

As a demonstration of the robustness of the complex suppression algorithm, the least-mean-squared error estimation was performed in the presence of background noise [i.e., $b(n)$ in Equation (7)] with no information signal present. These measurements, illustrated in Figure 13, were made using a synthetic seven-tone interference signal with linear FM (with constant amplitude, a fundamental frequency of 300 Hz with a frequency sweep of 50 Hz/s, and thus 350 Hz/s on the highest harmonic) in the presence of white Gaussian noise. Each successive harmonic is down by 6 dB from the previous. A 20-ms Hamming window, an 8-ms frame, and a 2048-point DFT were used. A longer window is used than in previous experiments to account for the lower frequencies present in the interference. The refined parameter estimation technique was found to be robust at INRs down to 0 dB. As in the single-tone case, the suppression ratio was determined by measuring the interference parameters in the presence of noise, suppressing the original interference signal (without noise present), and then comparing the power in the interference signal before and after suppression. Although the suppression ratio drops as INR decreases, as with single-tone suppression, the perceived interference residual in noise is removed even at the low INR of 0 dB. Figure 13 also shows that

the complex multitone suppression is less than for the single-tone (one harmonic) counterpart; the difference for a large range of INRs is nearly constant, an observation that leads one to consider the sidelobes from neighboring tones as an additional noise source.

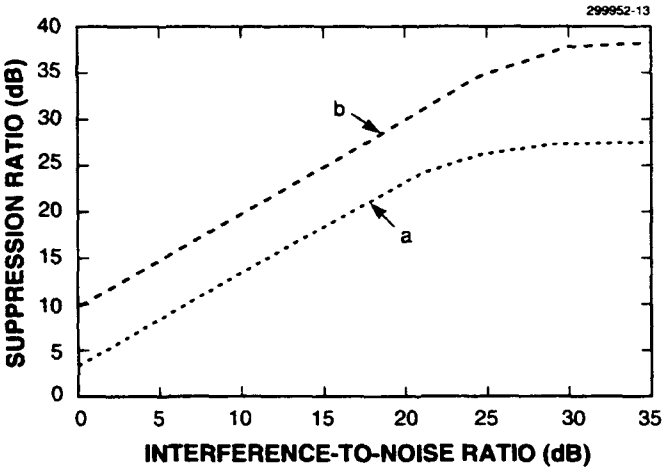


Figure 13. Suppression ratio as a function of INR: complex suppression on (a) multi- and (b) single tone (single harmonic signal).

4.3 Examples

Complex suppression provides a substantial reduction of the interference signal with the perceptual character of the information signal approximately preserved. A number of multitone examples are illustrated: a reference synthetic multitone and real signals, including acoustic signals from biologics, rubbing ice plates, and a siren disturbance. The examples illustrate both the features and the limitations of the approach. A 25-ms window, 10-ms frame, and 2048-point FFT were used. Spectral notch compensation was not applied to remove spectral bias.

4.3.1 Synthetic Multitone

In this example the synthetic interference signal consists of six harmonically related tones derived from a initial fundamental frequency of 250 Hz with linear frequency sweep of 50 Hz/s. The amplitude of the tones decreases by 6 dB as the harmonic frequency increases and is constant for each tone. The information signal is an acoustic signal from a closing stapler, and ISR is about 15 dB. Figure 14 gives a time-domain comparsion of the original and enhanced information signals, showing fidelity of the time structure in the reconstruction. In this case, about a 25-dB suppression

ratio was obtained so that the interference residual is about 10 dB below the information signal. Figure 15 gives a frequency domain view of a similar comparsion but with a white background noise added at a 20-dB INR. (The information signal is a closing stapler.) As expected, spectral nulls are seen in the background noise as a result of the spectral bias of the least-squares parameter estimator.

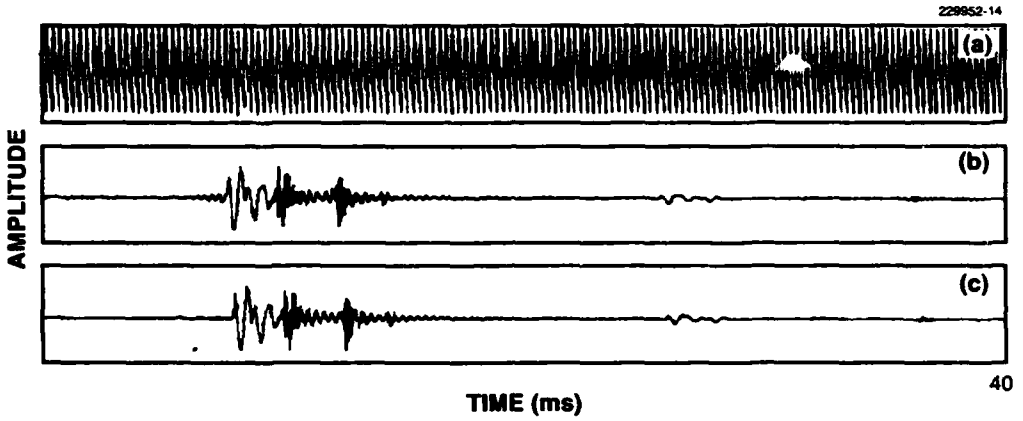


Figure 14. Multitone interference suppression: (a) interference with closing stapler, (b) processed, and (c) original closing stapler.

4.3.2 Biologics

Figure 16 illustrates an example of a (six-component quasi-harmonic) whale cry in an ocean background to which is added an information signal (the closing stapler). Six interfering frequencies were sought with harmonic guidance.[13] Harmonic interference is essentially removed, its perceptual character preserved, and as expected, spectral notches are placed at the peak locations of the interference. A second example of a biologic, illustrated in Figure 17, is the bark of a ringed seal to which is added the closing stapler. The interfering bark is characterized by rapidly and periodically varying AM. Although other bark harmonics are observed, the bark is dominated by its first harmonic, and in the vicinity of the peak amplitude of this component, the FM is roughly linear. The bark is the loudest tonal signal among the other interfering ocean and biologic tonal signals in the data. When only one frequency component is sought, the least-squares estimator

[13]The pitch estimation algorithm used in these examples is a derivative of a technique derived originally in the speech context based on a sine-wave representation of a signal [13].

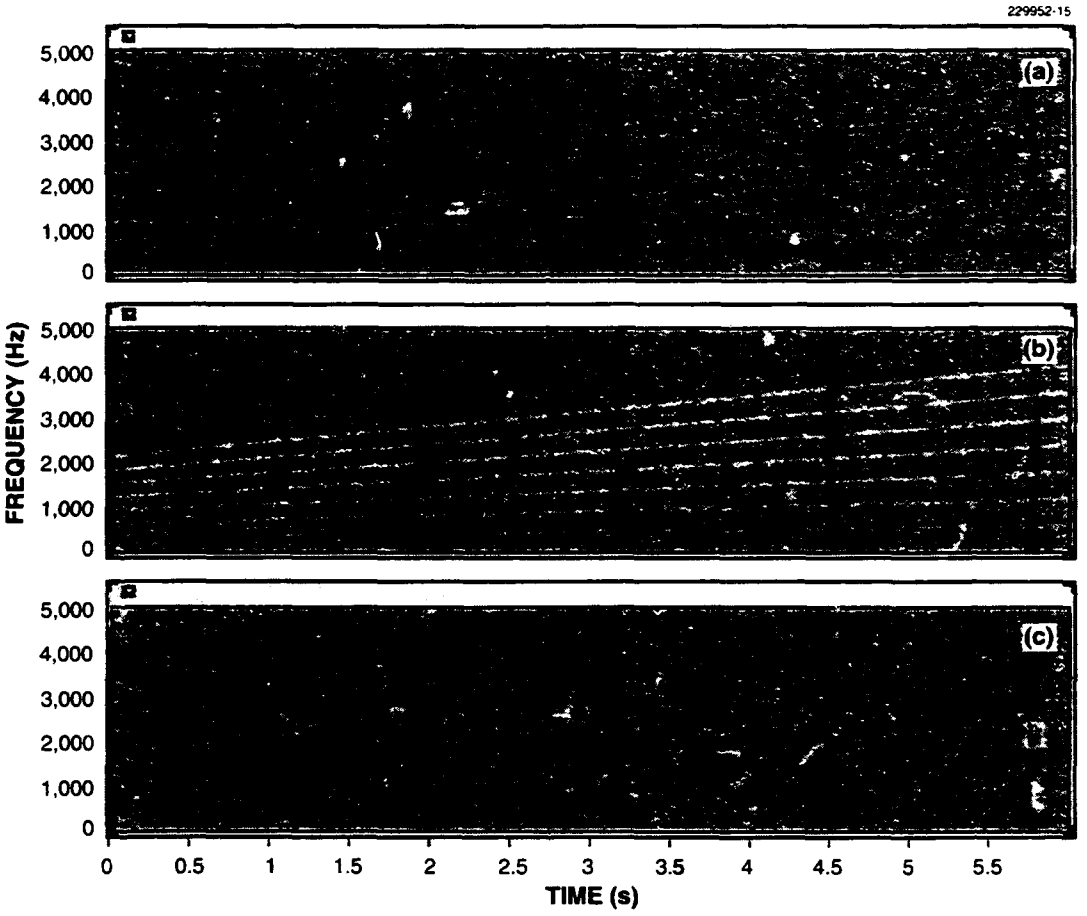


Figure 15. Multitone interference suppression in noise: (a) spectrogram of interference with closing stapler and background noise, (b) processed, and (c) spectrogram of original closing stapler in noise.

readily tracks the dominating first harmonic, thus effectively removing the bark without altering the background or the character of the information signal.

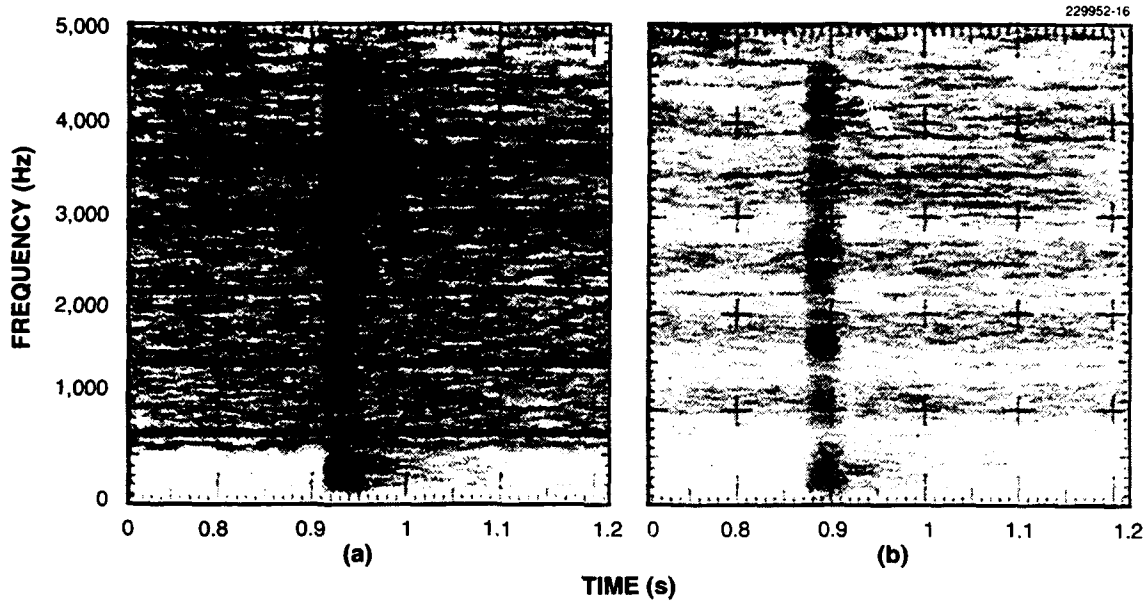


Figure 16. Suppression of interfering whale cry: spectrogram of (a) biologic signal with slamming book and (b) processed.

4.3.3 Slipping Ice Plates

Another signal that can interfere with underwater exploration is the sound generated from slipping or rubbing ice plates. Such signals, although often quasi-harmonic, may consist of more than one harmonic set due to slippage and cracking of the ice. These signals may consist of rapidly varying and discontinuous FM, making their suppression particularly difficult. An example of suppression of an acoustic signal emitted from ice is shown in Figure 18, where the ice signal comprises a slowly varying harmonic FM. (The closing stapler was added to the interference.) The first four harmonically related tones are removed. The harmonic nature of the interference allows the use of a pitch contour as a guide in suppression of the desired four tones. In this example the frequency guides are necessary to avoid unwanted local minima in the least-squares error minimization due to the presence of other tonal background signals.

A second example of suppression of a more complex ice signal is shown in Figure 19(a), (b). The acoustic signal from the ice slippage in this case consists of two harmonic sets that

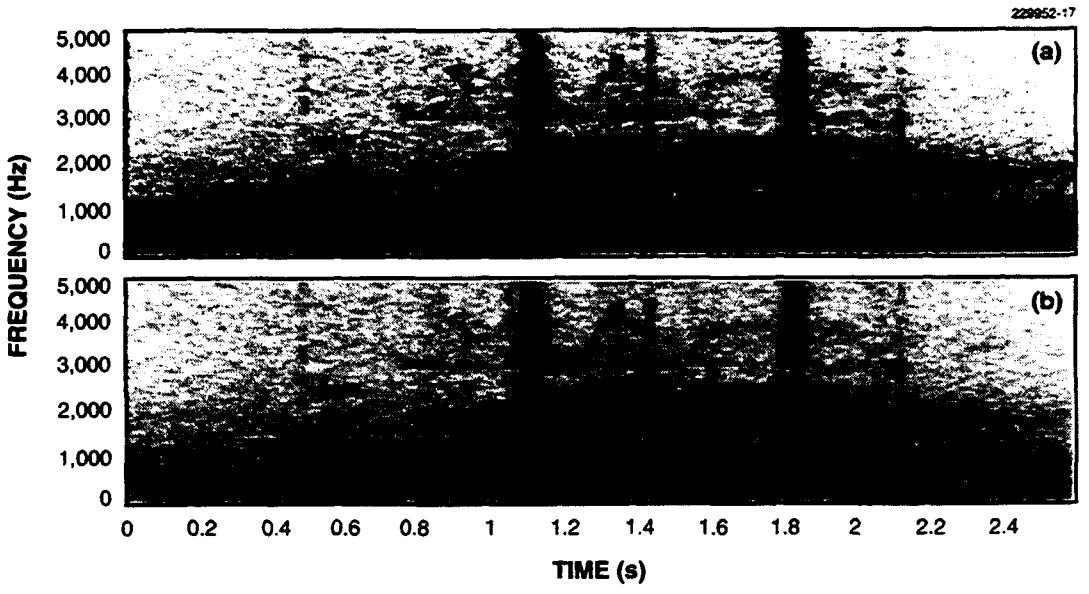


Figure 17. Suppression of interfering seal bark: spectrogram of (a) biologic signal with closing stapler and (b) processed.

intersect. Two sets of harmonic frequency guides are shown in Figure 19(c); the fundamental frequencies of each set were obtained, in part, manually and, in part, by the sine-wave-based pitch estimator. Although suppression is generally effective, residual is observed in the regions where the two harmonic sets intersect, at which point the single-tone linear AM-FM model is violated. Another condition in which the interference model does not hold is illustrated in Figure 20. In this case, the frequency tracks are characterized by sudden discontinuities in frequency (or pitch) that result in perceived glitches in the enhanced signal. (This example is further explored in Appendix E.)

4.3.4 Siren Disturbance

As a final multitone example, a synthetic siren was generated using frequency characteristics measured from an recorded siren.[14] The fundamental frequency trajectory of the synthetic siren was generated by fitting a fourth-order polynomial to measured points of the trajectory (of the siren's fundamental frequency), providing a frequency function for one cycle of the siren's frequency trajectory. This frequency function is then repeated periodically. The fourth-order polynomial was

[14]An actual siren was not used because of the unavailability of an uncorrupted recorded siren with a desirable information signal.

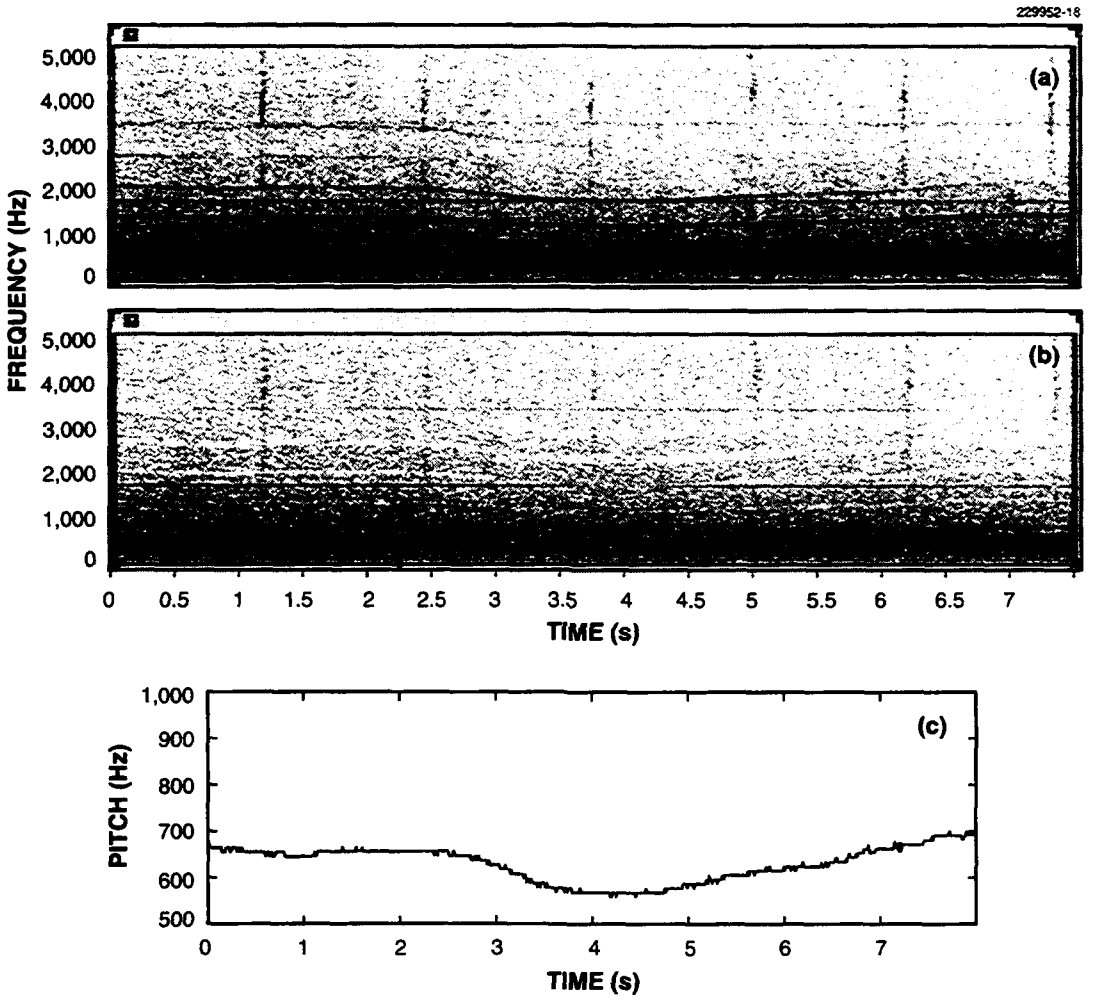


Figure 18. Suppression of high-pitch ice: spectrogram of (a) ice with closing stapler, (b) processed, and (c) pitch contour.

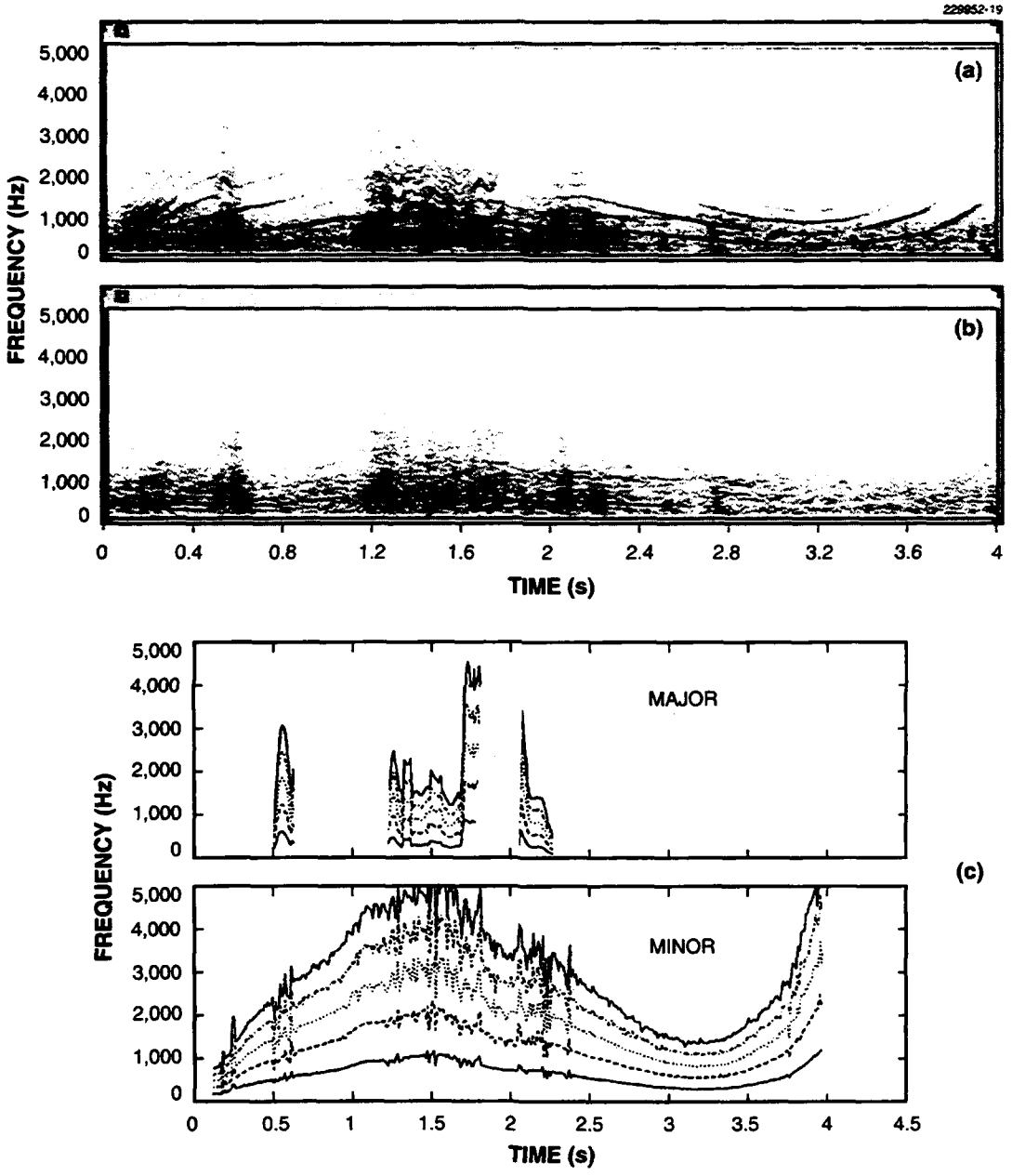


Figure 19. Suppression of two-pitch ice: spectrogram of (a) ice, (b) processed, and (c) harmonic guide contours.

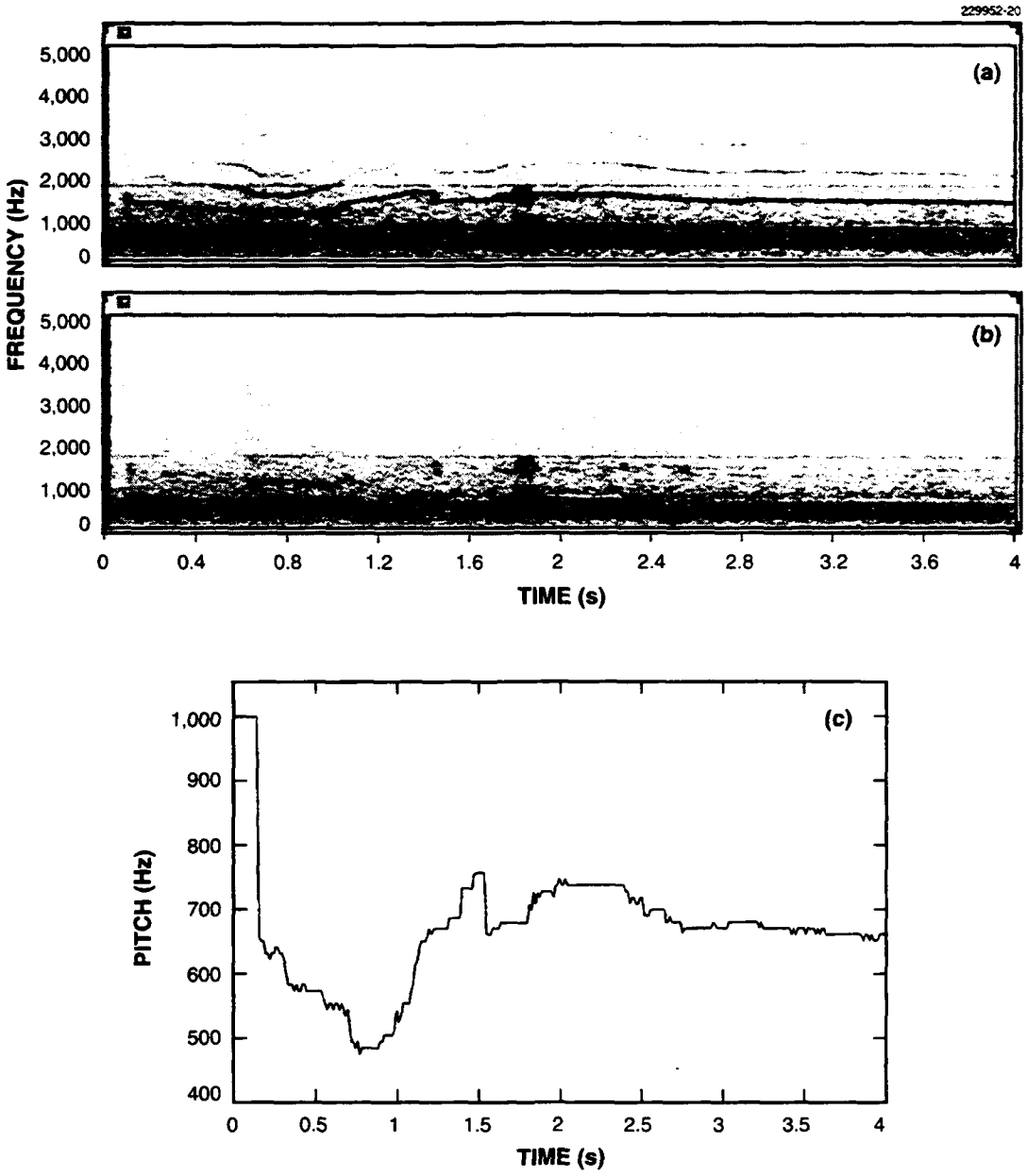


Figure 20. Suppression of erratic ice: spectrogram of (a) ice, (b) processed, and (c) pitch contour.

multiplied by two and by three to generate frequency trajectories for the second and third harmonics of the siren. The phase functions are generated by integrating the instantaneous frequency functions.

The synthetic siren is added to a voice (i.e., the information signal) with roughly a 25-dB ISR, making the voice barely audible. After suppression, the voice is significantly enhanced (see Figure 21). A small background residual from the siren remains, except for short "bleeps" where the discontinuity in frequency derivative[15] of the synthetic siren occurs and which cannot be accounted for by the linear-FM model under study (see Appendix E). The vertical striations that are observed in the spectrograms of both the original and processed signals result from this discontinuity.

4.4 Discussion

As with single-tone, multitone suppression suffers from (multiple) spectral nulls in a noise background. The multiplicity of these nulls makes the problem of background preservation particularly important from an aural perspective. A compensation filter can be derived as a generalization to the filter in Equation (21). Another challenge is the presence of more than one harmonic set or the presence of multiple aharmonically related frequencies created by a nonlinear medium. Improved multisignal pitch estimation and nonlinear prediction of such frequencies, to create frequency guides, will be useful. Another issue is the selection of a window when analyzing multiple tones. Ideally, the window should be long for low- and short for high-frequency tonal components or transients, requiring a multiresolution analysis/synthesis. Finally, the limitation of the suppression algorithm for discontinuous frequency (and frequency derivative) trajectories was observed; possible solutions to this problem are discussed in Appendix E.

[15]Both phase and frequency functions are continuous; however, because the frequency trajectory is derived by concatenating one function over successive periods, its derivative is discontinuous at the end of each period.

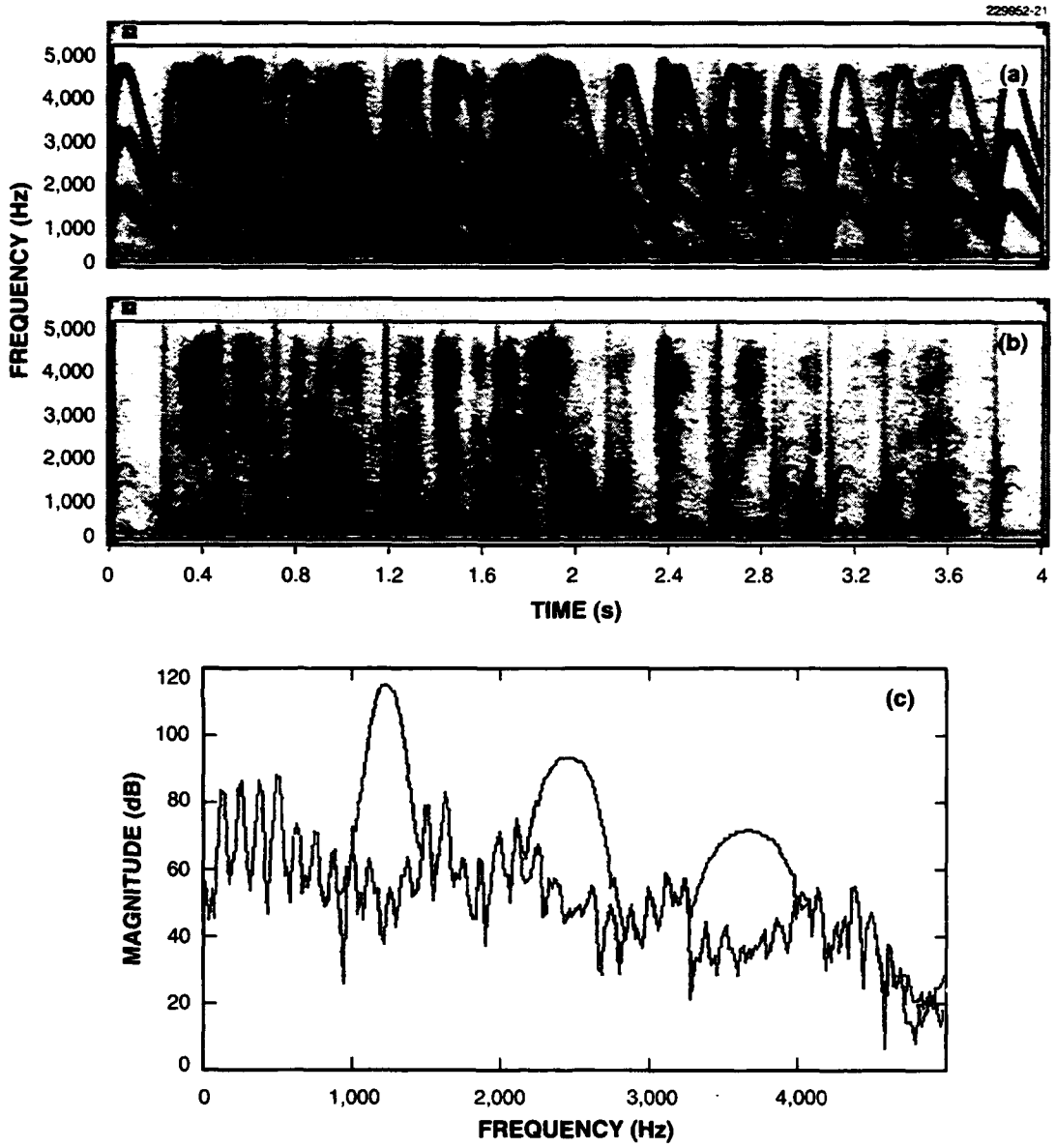


Figure 21. Suppression of complex siren disturbance: spectrogram of (a) original, (b) processed, and (c) spectral slices of original and processed signals overlaid.

5. MULTICHANNEL BEAMFORMING

In a multichannel (spatial array) environment, beamforming can be used to suppress interference coming from directions other than a desired information signal. Certain interference, however, (e.g., a blaring siren) can be sufficiently high to leak through the beamformer sidelobes and dominate the desired signal. Furthermore, putting a deep null in the direction of the interference signal can distort the main beam; the approach fails completely when the interfering and information signals lie in the same direction. This section explores use of the complex suppression algorithm on multichannels prior to beamforming to enhance beamformer performance.[16]

5.1 Problem Formulation

The multichannel suppression problem can be formulated as follows. Denote the received signal on the kth channel by $r_k(n)$. Then

$$r_k(n) = d(n + k\Delta_d) + i(n + k\Delta_i) \quad , \tag{30}$$

where $d(n)$ and $i(n)$ are the desired information and the interference signals, respectively, at a reference sensor, and where the delay terms Δ_d and Δ_i are a function of the arrival directions of the desired and interference signals. The enhanced signal after suppression on each channel is then given by

$$q_k(n) = S[r_k(n)] \quad , \tag{31}$$

where S denotes the suppression operator and can be written as

$$q_k(n) = \hat{d}_k(n) + e_k(n) \quad , \tag{32}$$

where $\hat{d}_k(n)$ is an estimate of the desired signal and $e_k(n)$ is the interference residual at the kth channel. A simple delay-and-sum beamformer is given by

$$q(n, \Delta) = \sum_k q_k(n - k\Delta) \quad , \tag{33}$$

---

[16]Suppression can also be performed after beamforming. In this case the interference residual may be on the order of the information signal.

where Δ is the interchannel delay, which in relation to the array spacing determines the look direction of the array.

In applying interference suppression to each channel prior to beamforming, it is important that the underlying signal of interest $d(n)$ not suffer from a phase distortion, which can degrade beamformer performance as measured through its array gain [14]. For example, if on each channel the distortion on $d(n)$ is simply a random phase δ_k that results from suppression, then the result is

$$\tilde{q}(n, \Delta) = \sum_k d_k(n - k\Delta + \delta_k) + e_k(n - k\Delta) \quad . \tag{34}$$

An alternate problem that may degrade performance is a possible correlation in the interference residual across channels, and this correlation exhibits itself as artifacts after beamforming (i.e., beamforming enhances the residual). For example, if each residual were identical across channels, then Equation (33) becomes

$$q(n, \Delta) = \sum_k \hat{d}_k(n - k\Delta) + e(n - k\Delta) \quad , \tag{35}$$

where the residuals on each channel are related by a delay.

The following examples demonstrate that interference suppression does not degrade beamforming performance. On the contrary, beamforming with the multichannel preprocessing enhances the signal of interest while further reducing the interference.

5.2 Examples

Experiments were formulated with a simulated 16-channel linear array of elements. The interference and information signals were summed with delays corresponding to different angles of arrival. A background noise scenario was simulated by using 16 different white Gaussian noise sequences for background noise and adding to each the interference and desired signal. In the following examples, a 100-ms analysis window was used (because the interference FM is slowly varying), and overlap-add synthesis was performed, although sine-wave synthesis gives similar results.

5.2.1 Multitone Interference without Background Noise

A multitone signal with FM (250-Hz initial fundamental frequency with a 50-Hz/s linear sweep, six harmonics, and constant amplitude) is added to the acoustic signal from the closing stapler with an ISR of about 15 dB. The interchannel delay of the information signal is zero (for an angle of arrival of 90°) while the interference has an interchannel delay of 0.1 ms (the corresponding angle of arrival depending on the distance between array elements). Background noise is not present, and the 16 channels are summed with zero delay between channels to generate the beamformed

output in the direction of the information signal (the closing stapler). Figure 22 illustrates that beamforming enhances the signal fidelity over the single-channel suppression case.

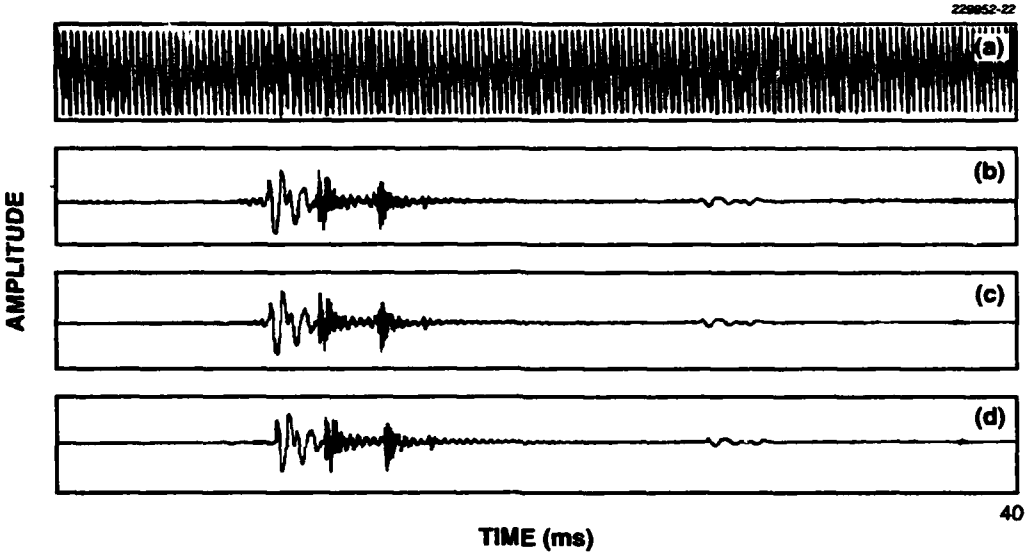


Figure 22. Interference suppression followed by beamforming: (a) interference with stapler, processed with (b) 1 channel, (c) 16 channels, and (d) original stapler.

5.2.2 Multitone Interference with Background Noise

The interchannel delay of both the interference and information signal is zero, corresponding to an angle of arrival of 90° between the array elements and these two signal components (an interchannel delay of zero).

In the first example the desired signal is a weak tone at 1000 Hz, and the interference is a synthetic, four-tone linear-FM chirp signal with a 700 Hz fundamental and a 50-Hz/s sweep frequency.[17] The interference suppression algorithm selects the strongest tone as the fundamental chirp frequency and uses this frequency and its three harmonics as suppression guides so that the tone of interest is not affected by suppression.[18] Figure 23(a) shows spectrograms of the weak tone

[17]A weak tone was specifically chosen for the information signal to provide a good test for the beamformer.

[18]It is assumed that the tone of interest has a lower power level than the fundamental of the interference signal.

in noise, the weak tone in noise with the synthetic linear-FM interference, and the processed version of the latter signal. (The tone is so weak that it cannot be seen in the spectrogram of a single channel.) The figure illustrates that the interference is suppressed without introducing artifacts, barring spectral nulls in the vicinity of the chirp interference. In Figure 23(b), spectrograms of the beamformer output for an interchannel delay of zero samples are shown for the weak tone in noise, the weak tone with interference in noise, and the processed version of the latter signal. The weak tone becomes visible as a result of beamforming, while the interference is effectively suppressed. Figure 23(c) shows spectra of the beamformer output with respect to interchannel delay (i.e., angle of arrival), where a delay of zero corresponds to the direction of the interference and signal of interest. These displays are shown for the weak tone in noise, the weak tone with interference in noise, and the processed version of the latter signal. These last displays were formed by averaging the magnitudes of four 1024-point DFTs of sequential segments of the beamformed output at each interchannel delay.

When interference is present, the tone of interest cannot be seen after beamforming because it is obscured by the sidelobes of the interference spectrum with the additional problem that the interference may be misconstrued as a signal of interest. With the application of interference suppression prior to beamforming, the interference is removed and the tone becomes visible. Figure 23(d) gives a different perspective of this performance, showing spectral slices of the beamformed output for an interchannel delay of zero and also showing that without preprocessing, the tone of interest is obscured by the interference but is easily detected when interference suppression is applied prior to beamforming.

The next example, illustrated in Figure 24, is identical to the previous example but with the tone of interest at a higher power level. The tone is visible in the spectrogram of a single channel prior to beamforming. In this case the tone of interest has enough power so that in Figure 24(c) its spatial sidelobes can be seen. These sidelobes are not disturbed by interference suppression, and as before no artifacts are introduced.

5.3 Discussion

The effectiveness of multichannel suppression prior to beamforming has been demonstrated. It was shown that the interference residual neither introduces artifacts in the beamformed output nor degrades the beamformed information signal. On the contrary, beamformer performance improved. An implication is that phase distortion (dispersion) in the information signal or correlation in the residual across channels is negligible.

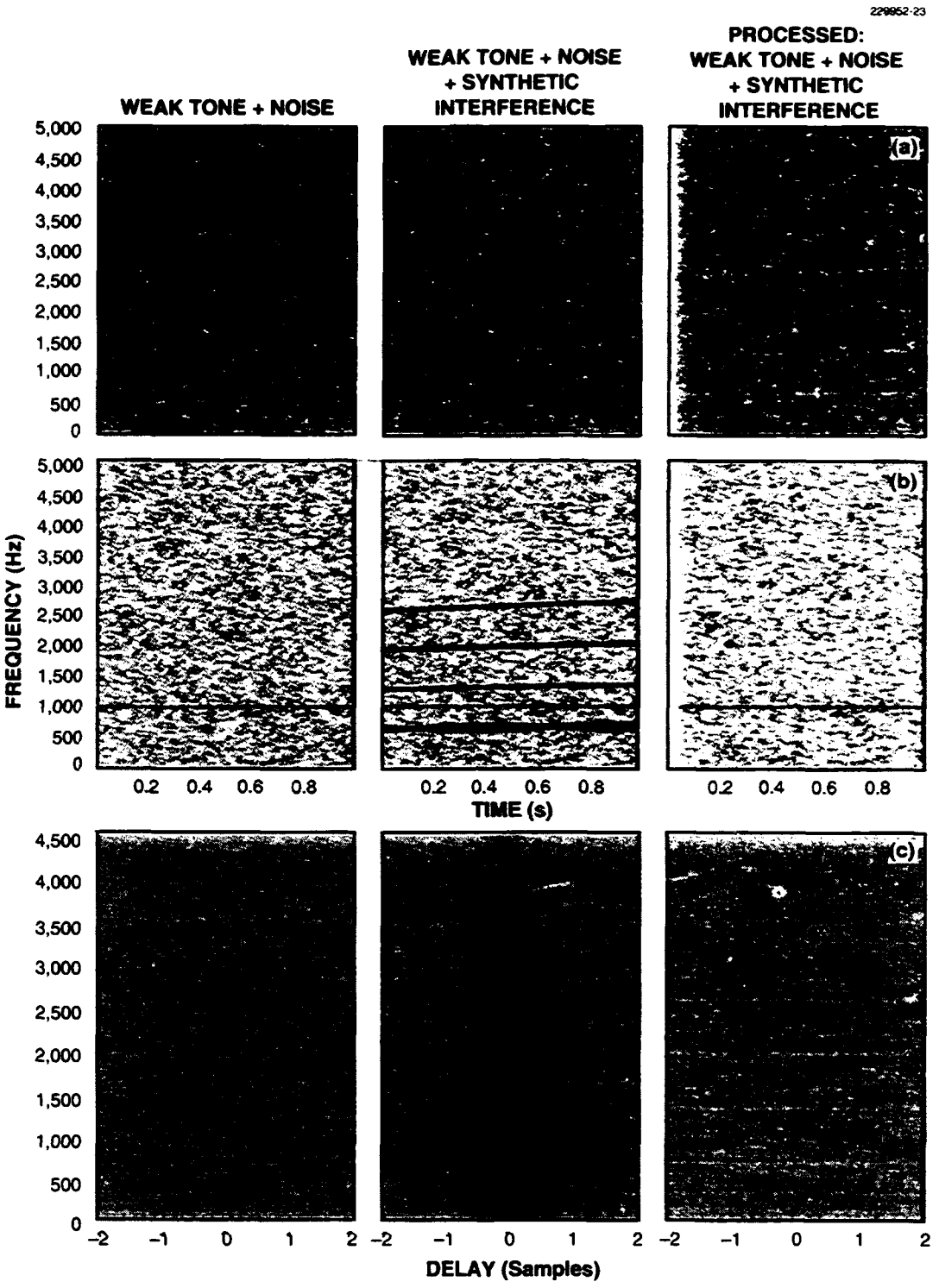


Figure 23. Synthetic example, case 1—weak synthetic tone and synthetic linear-FM interference with harmonics and broadband background noise: spectrograms of (a) single channel, (b) beamformed output at an interchannel delay of zero, (c) spectrum of beamformed signal versus interchannel delay. [Figure 23(d) follows.]

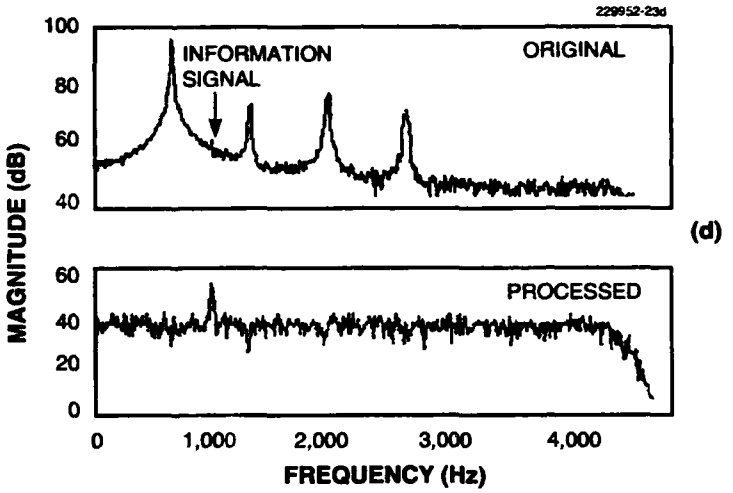


Figure 23 (Continued). Synthetic example, case 1: (d) spectral slices of beamformed output (at an interchannel delay of zero) for signal with and without preprocessing.

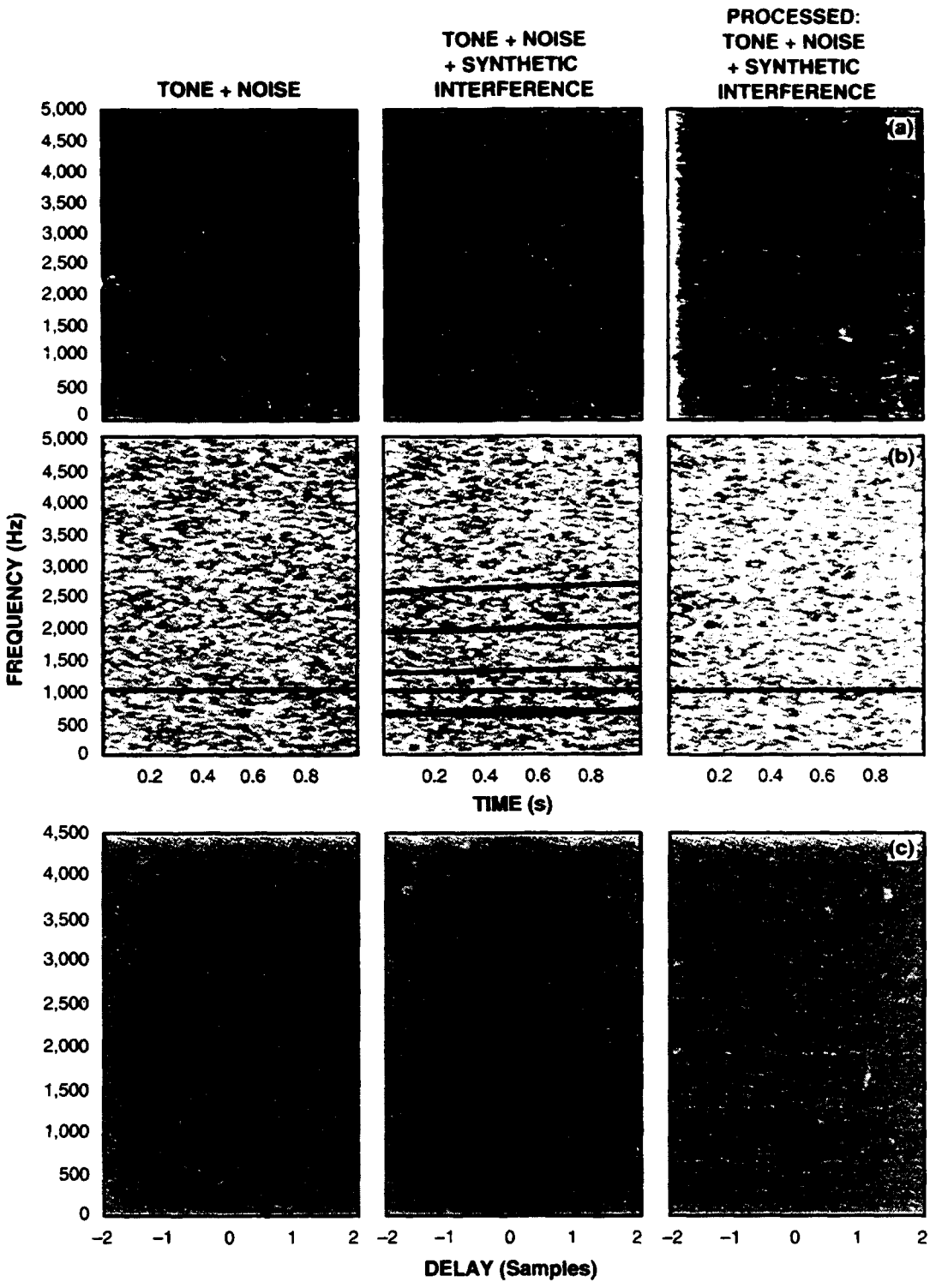


Figure 24. Synthetic example, case 2—strong synthetic tone and synthetic linear-FM interference with harmonics and broadband background noise: spectrogram of (a) single channel, (b) beamformed output at an interchannel delay of zero, and (c) spectrum of beamformed signal versus interchannel delay.

6. SLOW-MOTION AUDIO REPLAY

In slow-motion audio replay, the magnitude, frequency, and phase of the sine-wave components are modified to expand the time scale of a signal without changing its frequency characteristic (see Figure 1). This modification can be performed jointly while suppressing AM-FM tonal interference, yet also preserving a random background.

6.1 The Algorithm

Consider a time-scale expansion by a factor of β. By time-expanding the sine-wave frequency tracks, i.e., $\omega(\beta t, k) = \dot{\theta}(\beta t, k)$, the instantaneous frequency locations and magnitudes are preserved while modifying their rate of change in time. Because $d/dt[\theta(t\beta, k)/\beta] = \omega(\beta t, k)$, this modification can be represented by

$$\tilde{s}(t) = \sum_{k=1}^{N} A(\beta t, k) \cos[\theta(\beta t, k)/\beta] \quad . \tag{36}$$

The discrete-time implementation of Equation (36) requires mapping the synthesis interpolation frame duration Q to βQ, and then sampling over this longer frame the modified cubic phase and linear amplitude functions derived for each sine-wave component.

An example of slow-motion audio replay applied to the closing stapler is illustrated in Figures 25(a) and (b), where a sequence of events are time expanded. Each component lingers over a longer duration than the original, the effect of which is greater perceived separability of the time events and a sharpening of the spectral resonances. In informal listening, the audibility of the stapler's rapidly changing sequence of events is enhanced.

6.2 Background Preservation

As with interference suppression, signal modification should be designed so that the character of the resulting background is not altered. For random backgrounds it was found that large time-scale expansion may result in synthesized sine waves being perceived as tones, thus destroying the noise-like character of the original background. The problem is that the long synthesis frames, resulting from a large factor β, impose a time correlation on the sine-wave amplitudes and phases that does not exist in the representation of the original background. To avoid this objectionable tonality, a method is being developed to decorrelate the sine-wave phases across successive frames.

The essence of the technique is to add a random element to each sine-wave phase prior to doing cubic phase interpolation in the synthesis stage. This perturbation, although decorrelating the background phases, also decorrelates the phases of the information signal. Consequently, adaptive procedures are being developed that add the phase perturbation only when the background is

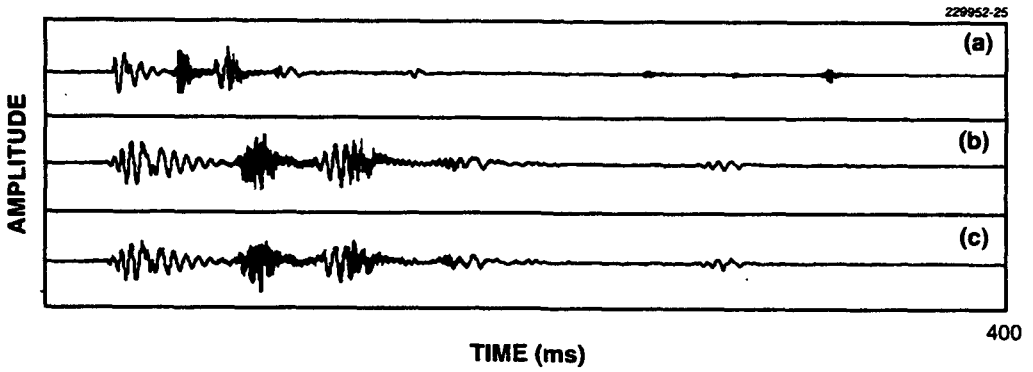


Figure 25. Example of slow-motion audio replay, stapler: (a) original, (b) after time-scale expansion by a factor of 2, and (c) after combined interference suppression and time-scale expansion by a factor of 2.

present (i.e., the information signal is not present). In one approach the modified phase for each frame m is given by

$$\tilde{\theta}(m,k) = \theta(m,k) + \epsilon(m,k) \quad , \tag{37}$$

where

$$\epsilon(m,k) = \pi\delta D(m,k) \tag{38}$$

with δ a random number falling uniformly in the interval $[-1,1]$, and $D(m,k)$ takes on the value zero when an information signal is present for the kth sine wave and one otherwise. This detection is performed by comparing the instantaneous energy in each band with a threshold derived from a running average energy. This approach and its derivatives have shown promise in preserving the background noise character while keeping the desirable properties of the time-scaled information signal. More extensive evaluation and alternative structures for both background preservation and improved temporal resolution of the sine-wave modification system are given in Quatieri, Dunn, McAulay, and Hanna [15].

6.3 Joint Modification and Suppression

The flexibility of the sine-wave signal representation allows signal modification to be performed jointly with interference suppression. From Equations (15) and (37), modification can be performed using the sine-wave amplitudes and phases to which interference suppression has been

applied. Figure 25(c) illustrates these joint operations in which the response of the closing stapler is corrupted by the AM-FM interfering tone used in Figure 5.

6.4 Discussion

An advantage of the sine-wave framework for suppression is its straightforward integration with signal modification schemes. A time-scale modification method was presented, but frequency modifications are also being considered. One approach to preserving the character of the time-scaled background was described. When applying interference suppression jointly with time-scale modification, the method of preserving background can be integrated with the preservation method developed in Section 3.5 for compensating a spectral null.

7. SUMMARY AND FUTURE WORK

A new approach to interference suppression has been developed to enhance the audibility of acoustic signals. The technique is applicable to single as well as multiple AM-FM tone suppression and is robust in complex backgrounds. Because the approach was developed in the framework of sine-wave analysis/synthesis, the suppression can be integrated with signal enhancement by slow-motion audio replay. This technique is one of a class of signal modifications being developed, including rapid audio scanning and sine-wave frequency manipulations. Random backgrounds are approximately preserved with either suppression or modification by appropriate estimation and manipulation of sine-wave components, a property that is essential for minimizing false detection of information signals. Finally, it was shown that interference suppression on multichannels prior to beamforming enhances beamformer performance. Although significant audibility gains were achieved, much remains to be accomplished. Important directions were discussed throughout the report; an overview of these future efforts is summarized next.

Selection of the Analysis Window: One unresolved area is the selection of an "optimal" window duration over which to perform suppression. Ideally, the window duration as well as other algorithm parameters should be tailored to the characteristics of the interference and information signals, for example, a slowly or rapidly varying FM, a sharp or gradual onset, and the number and orientation of tonal components. Although Appendix D formulates certain informal rules for this selection, a more rigorous approach awaits.

AM-FM Discontinuities: Related to the selection of the analysis window duration is the problem of discontinuity in AM and FM. In a real-world situation, the AM and FM (and their derivatives) of the interference signal may be characterized by abrupt changes, as in a pulsed siren or the sudden change of ice movement. These discontinuities introduced into the interfering signal violate the assumed model because the signal under the analysis window is not accurately modeled by a linear AM-FM signal, making both parameter estimation and interference subtraction prone to error and resulting in artifacts, which may be misinterpreted as information signals. One approach to reducing these effects is to adapt the analysis window to the interference by shifting the analysis window so that stationary regions of the interfering signal lie within its extent. One approach to selecting such regions is proposed in Appendix E.

Background Preservation: Another area for future work is the continued development of methods to reconstruct the background signal in regions where the suppression algorithm results in spectral nulls; such spectral reconstruction improves both aural and visual displays. Alternative methods of spectral extrapolation should be considered, such as white-noise driven synthesis and methods in the style of band-limited extrapolation. Another remaining problem is integrating detection (of the information signal) with determining the appropriate time-frequency extrapolation.

Suppression in Presence of Information Signal: A thorough evaluation of the algorithm in the presence of information signals has not been performed, only in the presence of noise. In

this context, it may be of interest to "close the loop": after suppression, subtract the information signal estimate from the received signal and then repeat the least-squares parameter estimation.

Computational Complexity: To make the algorithm feasible, it is necessary to reduce the complexity of the iterative technique that solves the least-squares parameter estimation problem and is the dominant computational burden within the suppression algorithm.

APPENDIX A
State of the Art

A.1 Estimation of Linear FM

A number of techniques exist for estimating parameters of an AM-FM tone with linear FM and constant amplitude. This overview provides a flavor of the state of the art. A more exhaustive tutorial can be found in Boashash [8].

Bello [16]: A maximum-likelihood method was proposed for estimating Doppler delay (chirp phase offset), Doppler (chirp center frequency), and Doppler rate (chirp frequency sweep rate) in radar returns. A calculation was made of the Cramer-Rao variance bounds for these estimates. Bello argued that under certain conditions, the maximum-likelihood estimate is close to the minimum variance (least-squares error) estimate. Abatzoglou [17] applied Newton's method to find the peak in the maximum-likelihood function. The procedure uses a coarse search followed by a fine search via Newton's method. With moderate frequency rates the method breaks down at about a 15-dB SNR, above which the Cramer-Rao bound is aproximately achieved.

Rao and Taylor [18]: A class of techniques estimates instantaneous frequency from the peak in numerous time-frequency distributions; for example, a coarse estimate of a time-varying frequency modulation can be obtained by tracking the peak in the STFTM. Rao and Taylor have shown that the peak in the Wigner-Ville time-frequency distribution results in an instantaneous frequency estimation that is optimal for linear FM signals with high to moderate SNR. This method, however, degrades significantly at low SNR.

Djuric and Kay [19]: An estimate of the chirp phase was made using a parametric representation of the phase (i.e., in terms of phase offset, frequency, and frequency sweep). Least-squares estimation of the phase is performed (with respect to the unknown three parameters) not on the waveform, but on the phase—an important distinction from earlier methods. This procedure implies that the phase must be unwrapped prior to estimation. Phase unwrapping puts constraints on the accuracy of the frequency rate estimate, especially in noise. Because a large frequency rate and large noise can result in rapid phase jumps greater than 2π, ambiguity in the unwrapping process can result. For moderate frequency rates, the procedure breaks down at an SNR of about 10 dB, above which the Cramer-Rao bound is approximately obtained.

A.2 FM Interference Rejection

This section describes a number of techniques for rejecting FM tonal interference. As with Section A.1, this overview provides a flavor of the state of the art and not an exhaustive review.

Widrow [20]: An adaptive finite-impulse response (FIR) notch filter was derived using the LMS algorithm, which adapts the FIR filter coefficients as a function of time. This method requires first, explicitly estimating the frequency of the interference (i.e., obtaining a reference frequency) and second, implementing a notch filter at that interference. The method is capable of tracking very

slowly varying frequency interferences (FM) by adaptively tuning the notch filter (with a center frequency at the reference frequency) and can be generalized to multiple frequencies [4].

Rao and Peng [5]: An infinite impulse response (IIR) tracking notch filter was derived to estimate the coefficients using a Gauss-Newton algorithm. This approach is porported to have greater efficiency than adaptive FIR notch filters. Approximate and simple closed-form results were derived for the tracking behavior of a second-order notch filter. In particular, for very slow frequency variations (FM) in the signal, the behavior of the adapted filter coefficients can be studied as a solution to a differential equation.

Wulich, Plotkin, Swamy [21]: The problem addressed is that of estimating the parameters of a sine (e.g., amplitude and phase offset) in the presence of a closely spaced FM interference with fast frequency modulation. A discrete-time differential equation is formulated as a notch filter for FM signals of an arbitrary modulating function. The coefficients of the differential equation are a function of the instantaneous frequency of the FM estimated using a phase locked loop. (A fixed notch filter is first applied to remove a desired sine signal under the assumption that its frequency is known—clearly not practical for wideband signals.) The system was demonstrated at a 30-dB FM INR. This method is claimed to be more effective than linear FIR or IIR adaptive filtering, and it can be improved by warping the time axis according to the instantaneous frequency estimate, resulting in a tone without FM. A notch is then applied to the constant-frequency tone in the new time axis and the inverse operation is applied to obtain the enhanced signal [6].

APPENDIX B
Complex Suppression with Coarse Estimation

This appendix compares interference parameter estimation using the Powell least-mean-squared (PLMS) algorithm [9,10] with the discrete Fourier transform (DFT) approach (using the maximum spectral value and associated parameter estimates as derived in Section 3.1).[19] The interference signal used in this experiment is a tone with a sinusoidally varying instantaneous frequency $\omega(t) = 1500 + 400\sin[2\pi(0.532)t]$ (so that the frequency comprises a center frequency of 1500 Hz with a swing of 400 Hz and a maximum slope of about 1500 Hz/s) and a sinusoidally varying amplitude $A(t) = 1 + 0.2\sin[2\pi(0.617)t]$ (so that the amplitude comprises a constant of unity with a swing of 0.2 and a maximum amplitude slope of about 0.6/s). Because these modulations were selected to avoid regularities in the waveform, and because the interference does not strictly follow our short-time linear assumption, it provides a good test of the suppression algorithm. Analysis parameters were a 10-ms Hamming window, a 4-ms frame, and a 2048-point DFT. As illustrated in Table B-1, the PLMS method is clearly preferred, yielding a higher suppression ratio whether using magnitude-only or complex subtraction. In addition, as expected the segmental SNR improves with the refined suppression; however, the comparison of the magnitude-only and complex suppression must be considered with care due to the presence of interference residual.

TABLE B-1

DFT versus PLMS Parameter Estimation

Interference Parameter Estimation Method	Coarse (DFT)		Refined (PLMS)	
Suppression Method	**Magnitude**	**Complex**	**Magnitude**	**Complex**
Suppression ratio	39.00 dB	18.70 dB	52.50 dB	38.80 dB
Segmental SNR	2.70 dB	–14.30 dB	3.10 dB	2.80 dB

[19]Table B-1 always uses the DFT approach for A_o so that the comparison entails estimating A_s, ω_o, ω_s, and ϕ_0 by either the DFT or PLMS approaches. Eliminating this fifth variable from the search in the PLMS approach was found to significantly reduce computational time. Moreover, the suppression gained by estimating A_o via the PLMS approach was marginal.

APPENDIX C
Least-Squares Estimation

In solving for the parameters of the AM-FM tonal interference model, the error function is defined

$$\epsilon = \sum_n (w(n)[r(n) - i(n)])^2 \quad , \tag{C.1}$$

where $w(n)$ is the analysis window, $r(n)$ is the measurement, and $i(n)$ is the interference. The error ϵ is minimized over the parameters of the model for $i(n)$ given by

$$i(n) = (A_o + A_s n)\cos[\omega_o n + \omega_s \frac{n^2}{2} + \phi_o] \quad , \tag{C.2}$$

where both the amplitude and frequency are modeled by a linear trajectory. The highly nonlinear problem of minimizing ϵ with five free parameters can be solved with various well-established iterative methods [8,10]. One possibility is to perform an exhaustive search over a plausible parameter range; having a coarse initial estimate of the parameters allows defining such a parmeter range. Although this approach is typically computationally intractable, it does provide insight into the error surface associated with Equation (C.2). For example, when the parameters A_o, A_s, and ω_s in Equation (C.2) are held fixed, and the frequency ω_o and phase offset ϕ_o vary around a coarse estimate (derived from the peak frequency), then the error surface (locally) is found to take on an approximate "quadratic bowl" shape. A similar property was found when varying the frequency sweep ω_s and phase offset ϕ_o, while holding the remaining parameters fixed.

This observation motivates an iterative gradient descent procedure for minimization [20]. A problem arises, however, in this approach due to the need of computing derivatives (an intensive operation) and a feedback gain factor that must guarantee stability of the iterative descent under a variety of conditions. An alternative method is the Powell iterative method [9,10], which was selected for its computational ease and relatively rapid convergence; it requires neither the direct computation of derivatives nor a gain factor.

For the single-tone case, the starting point in the Powell method uses the coarse parameter estimates derived in Section 3.1. With this starting point, the iteration converges rather quickly (typically 5 with a maximum of about 20 iterations) to the desired local mimimum. As a demonstration of the robustness of the algorithm, the parameter accuracy of the least-mean-squared error estimation in the presence of background noise [(i.e., $b(n)$ in Equation (7)] is shown in Figure C-1. These measurements were made by comparing the known parameters of a synthetic interference signal (a linear FM sweep with constant amplitude, ω_0 = 1000 Hz, and ω_s=1000 Hz/s) with the parameters as measured in the presence of white Gaussian noise. A 10-ms Hamming window, a 4-ms

frame, and a 2048-point DFT were used. As seen in Figure C-1, the iterative least-squares method breaks down at an SNR of about 0 dB (i.e., where the knee in the curve occurs). With multiple tones, again using initial coarse estimates derived in Section 3.1, the Powell method was found to have properties similar to the single-tone case when each AM-FM tone is removed independently.

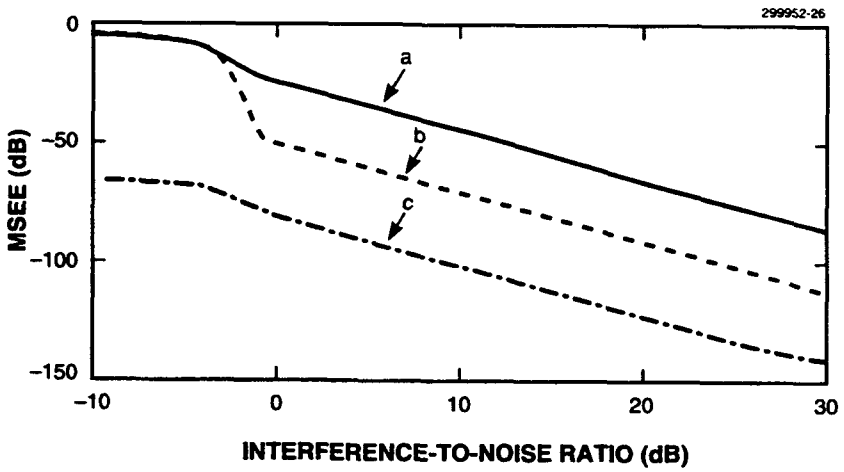


Figure C-1. Mean-squared estimation error (MSEE) versus INR for (a) ϕ_o, (b) ω_o, and (c) ω_s.

APPENDIX D
Analysis Window Selection

The purpose of this appendix is to give a flavor for the considerations required in the selection of the the analysis window length used in the suppression algorithms. This study was performed in the context of the overlap-add framework because window selection does not effect the reconstruction of the information signal (the overlap-add analysis/synthesis being an identity system). Similar considerations will hold for the sine-wave framework with respect to suppression; however, unlike the overlap-add framework, the window length must be considered with respect to the reconstruction of the information signal because as window length increases, time resolution decreases.

The goal of this study is to select a window duration that maximizes performance of the suppression algorithm while minimizing artifacts that might be introduced into the background. A measure of power removed was defined as the ratio of the average power in the signal to the average power in the processed signal. In numerous examples with this measure, it was observed that the power removed *decreases* as the length of the analysis window increases. Spectrogram analysis of the processed signals, however, revealed that interference suppression and preservation of the background spectrum are generally improved when the analysis window length *increases*.

To help isolate the cause of this apparent discrepancy, a controlled experiment was designed. A synthetic linear-FM interference signal was generated, comprising a chirp and four harmonics (650-Hz fundamental frequency with a 40-Hz/s linear sweep) with constant amplitude. White Gaussian noise was added at a 10-dB interference-to-noise level.[20] Figure D-1 shows that the power removed by the suppression algorithm decreases as the analysis window duration increases. The apparent contradiction is resolved by observing that the power removed from the received signal is due not only to the interference, but also to the background noise in the neighborhood of the frequency of the interference. The frequency band over which the noise is removed increases as the length of the analysis window decreases, which is consistent with the parameter estimator in Section 3 yielding a biased estimate of the background spectrum, forcing it to zero in the vicinity of the interfering chirp frequency. The suppression algorithm thus nulls the spectrum of the received signal in this region. As the duration of the analysis window decreases, the region over which the spectrum is nulled may increase; this nulling may be exacerbated by the accuracy of the interference parameter estimates decreasing as the window length decreases.

A third experiment was performed to verify this observation. The parameters of the synthetic interference were estimated in the presence of the white Gaussian noise background, and an estimate of the interference was reconstructed from the parameter estimates. The estimate of the interference

[20]Interference-to-noise level is defined as the ratio of the power in the interference to the power in the noise.

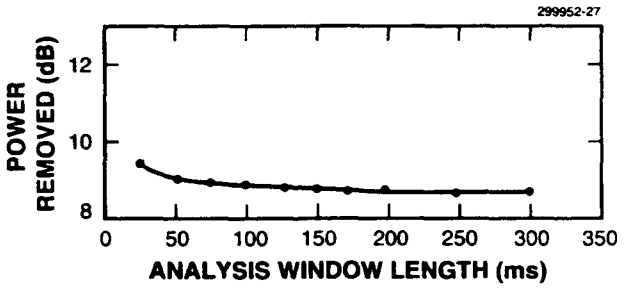


Figure D-1. Power removed from a synthetic signal with respect to analysis window length. The signal was a synthetic linear-FM with four harmonics and a white Gaussian noise background.

was then subtracted from the synthetic interference (without the noise present) to form a residual signal. A measure of interference suppression was defined as the ratio of the power in this residual signal to the power in the synthetic interference (the earlier defined suppression ratio). Figure D-2 is a plot of interference suppression versus analysis window length, demonstrating that interference suppression does indeed increase as the analysis window length increases.

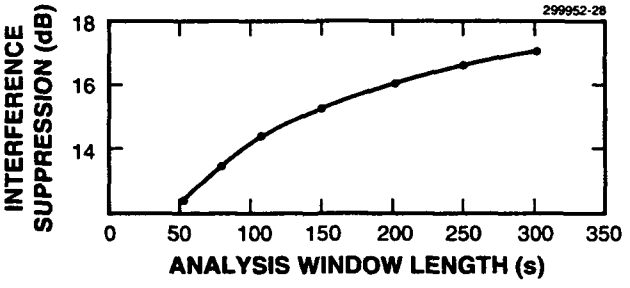


Figure D-2. Suppression of a linear-FM interference with four harmonics and a white Gaussian noise background with respect to analysis window length.

These experiments indicate that the "optimal" analysis window is the longest possible. This selection achieves a maximum degree of suppression and reduces background artifacts by decreasing the width of the spectral nulls; however, one must also consider that a long analysis window may lead to a data segment that violates the current linear-FM model. For rapidly varying FM, as well as for signals with abrupt onsets and offsets, the actual interference only approximately matches

the model, and this approximation improves if the analysis window is shorter. Consequently, a Hamming window of duration in the range 5 to 100 ms was generally chosen. The selection is a function of the characteristics of the specific data class.

APPENDIX E
Tracking Abrupt Frequency Changes

To account for onsets and offsets of the interference signal, as well as rapid variations in the AM and FM (which violate the linear AM and FM model), the analysis window duration should be made adaptive. One approach to achieve this adaptivity is first to track these changes and then shift the analysis window to encompass a quasi-stationary region of the interference. A new method for tracking such changes based on an "instantaneous" *energy operator* proposed by Teager [22–25] is being developed. The operator, representing the energy of a simple harmonic oscillator and originally developed in continuous time, has a discrete-time counterpart; a function of this discrete operator yields an estimate of the AM and FM of a signal using five time samples and thus has excellent time resolution. An example of the time resolution of the algorithm is illustrated in Figure E-1, where an abrupt change in a sine-wave frequency is tracked to within a few samples.

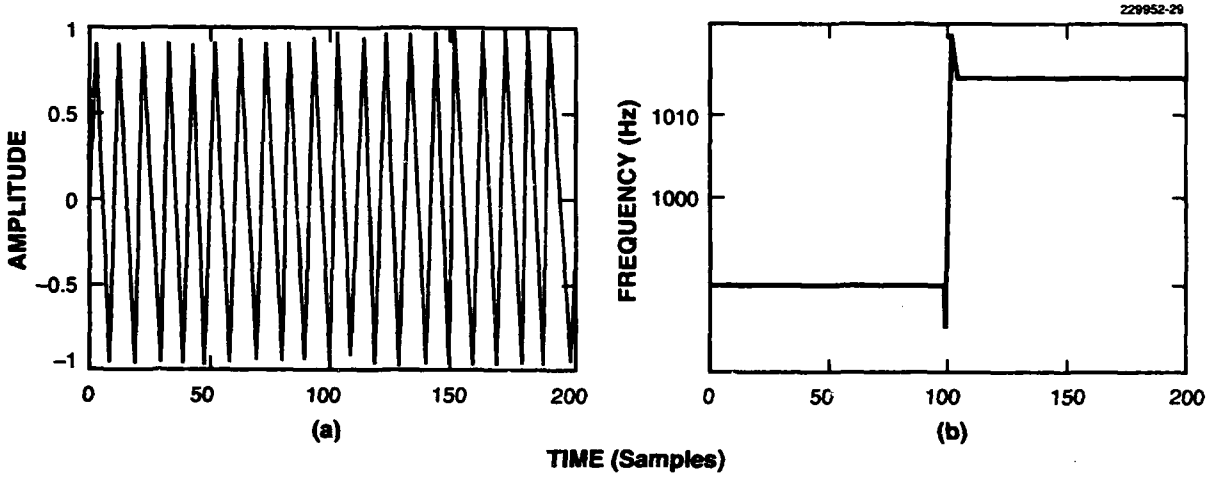


Figure E-1. Instantaneous frequency tracking using Teager operator: (a) waveform and (b) FM estimate.

This frequency tracker can be applied to the siren and erratic ice interference investigated in Section 4, both of which are characterized by abrubt change in FM. Figure E-2 shows the FM estimate of the first harmonic of the siren in the region of abrupt change as measured by the new operator; the frequency trajectory is characterized by a repeated discontinuity in its derivative, which violates the assumed model. Figure E-3 shows evidence of rapid frequency change of the second harmonic (obtained by bandpass filtering) in the erratic ice example of Section 4.3. In this

case it appears that the frequency change is not only abrupt at a specific time instant, but exhibits rapid oscillatory behavior prior to the change.[21] Given that the abrupt change may correspond to the slippage of two rubbing ice plates, the oscillatory frequency behavior may be a result of tension between the plates prior to the slippage. This rapid oscillation in FM (roughly four or five cycles over the duration of a 10-ms analysis window) violates the linear-FM model and may explain the increase in interference residual observed in the region of the abrupt frequency change.

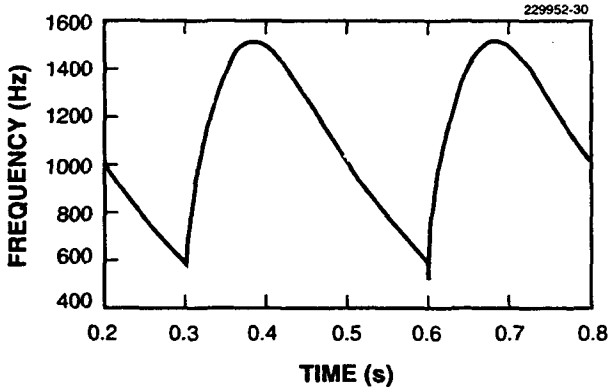


Figure E-2. Instantaneous frequency of first harmonic of siren using Teager operator.

[21] A more rigorous development of this approach requires showing that the observed frequency variations are not significantly influenced by the background noise.

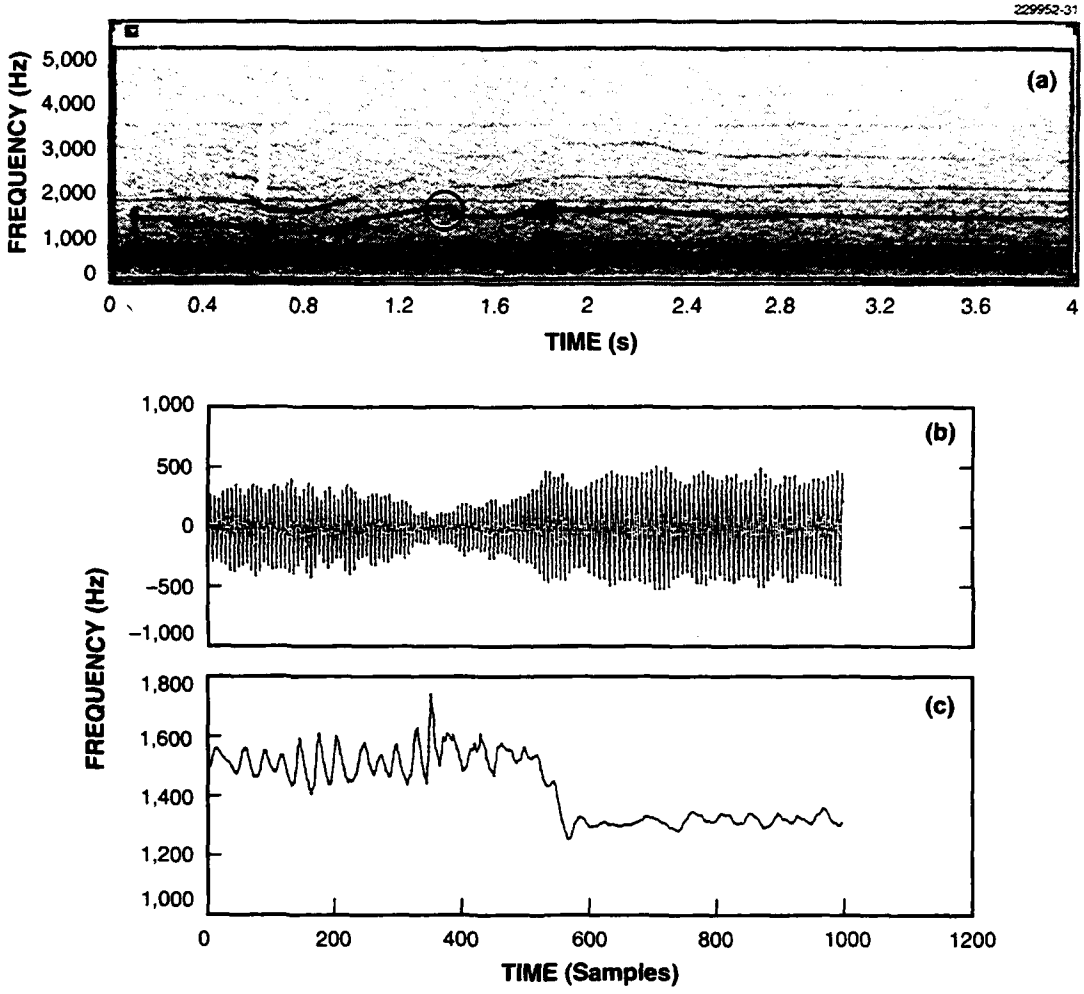


Figure E-3. Measuring abrupt frequency changes in ice using Teager operator: (a) spectrogram of ice, (b) bandpass-filtered second harmonic, and (c) FM estimate.

REFERENCES

1. R.J. McAulay and T.F. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," *IEEE Trans. Acoust. Speech Signal Process.*, **Vol. 34**, 744–754 (August 1986).

2. T.F. Quatieri and R.J. McAulay, "Sp formations based on a sinusoidal representation," *IEEE Trans. Acoust. Speech Si ess.*, **Vol. 34**,1449–1464 (December 1986).

3. T.F. Quatieri, R.B. Dunn, R.J. McAulay, T.E. Hanna, "Underwater signal enhancement using a sine-wave representation," *IEEE OCEANS '92 Conference*, Newport, R.I., 449–454 (26–29 October 1992).

4. J.R. Glover, "Adaptive noise cancelling applied to sinusoidal interference," *IEEE Trans. Acoust. Speech Signal Process.*, **Vol. 25**, 484–491 (December 1977).

5. B.D. Rao and R. Peng, "Tracking characteristics of the constrained IIR adaptive notch filter," *IEEE Trans. Acoust. Speech Signal Process.*, **Vol. 36**, 1466–1479 (September 1988).

6. D. Wulich, E.I. Plotkin, and M.N.S. Swamy, "Synthesis of discrete time-varying null filters for frequency-varying signals using the time-warping technique," *IEEE Trans. Circuits Syst.*, **Vol. 37**, 977–990 (August 1990).

7. H.S. Nawab, T.F. Quatieri, "Short-Time Fourier Transform" in *Advanced Signal Processing*, Englewood Cliffs, N.J.: Prentice-Hall (1988).

8. B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal," *Proc. IEEE*, **Vol 80**, 519–568 (April 1992).

9. R.P. Brent, *Algorithms for Minimization without Derivatives*, Englewood Cliffs, N.J.: Prentice-Hall (1973).

10. W.H. Press, B.P Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, Cambridge, England (1991).

11. S.M. Kay, *Modern Spectral Estimation*, Englewood Cliffs, N.J.: Prentice-Hall (1987).

12. A.V. Oppenheim and R.W. Schafer, *Digital Signal Processing*, Englewood Cliffs, N.J.: Prentice-Hall (1975).

13. R.J. McAulay and T.F. Quatieri, "Pitch estimation and voicing detection based on a sinusoidal model," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Albuquerque, N.M., 249–252 (3–6 April 1990).

14. W.S. Burdic, *Underwater Acoustic System Analysis*, Englewood Cliffs, N.J.: Prentice-Hall (1984).

15. T.F. Quatieri, R.B. Dunn, R.J. McAulay, T.E. Hanna, "Time-Scale Modification of Complex Acoustic Signals in Noise," Lexington, Mass.: MIT Lincoln Laboratory, Technical Rep. 990 (4 February 1994).

REFERENCES
(Continued)

16. P. Bello, "Joint estimation of delay, Doppler, and Doppler rate," *IRE Trans. Inf. Theory*, 330–341 (June 1960).

17. T.J. Abatzoglou, "Fast maximum likelihood joint estimation of frequency and frequency rate," *IEEE Trans. Aerosp. Electron. Syst.*, **AES-22**, 708–715 (November 1986).

18. B.D. Rao and F. Taylor, "Estimation of the instantaneous frequency using the discrete Wigner distribution," *Electron. Lett.*, **Vol. 26** 246–248 (1990).

19. P.M. Djuric and S.M. Kay, "Parameter estimation of chirp signals," *IEEE Trans. Acoust. Speech Signal Process.*, **Vol. 38**, 2118–2126 (December 1990).

20. B. Widrow and S.D. Starns, *Adaptive Signal Processing*, Englewood Cliffs, N.J.: Prentice-Hall (1985).

21. D. Wulich, E.I. Plotkin, and M.N.S. Swamy, "Discrete time-varying filter and PLL for synchronous estimation of parameters of a sine signal corrupted by a closely spaced FM interference," *Signal Process.*, **Vol. 21**, 183–194 (October 1990).

22. J.F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Albuquerque, N.M., 381–384 (3–6 April 1990).

23. P. Maragos, J.F. Kaiser, and T.F. Quatieri "On separating amplitude and frequency modulations using energy operators," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, San Francisco, Calif., II-2–II-4 (23–26 March 1992).

24. T.F. Quatieri, J.F. Kaiser, P. Maragos, "Transient detection in AM-FM background using an energy operator," Underwater Signal Processing Workshop, University of Rhode Island, Kingston (9–11 October 1991).

25. R.B. Dunn, T.F. Quatieri, and J.F. Kaiser, "Detection of transient signals using the energy operator," *Proc. IEEE Int. Conf. Acoust. Speech Signal Proc.*, Minneapolis, Minn., III-145–III-148 (27–30 April 1993).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	17 May 1994	Technical Report

4. TITLE AND SUBTITLE
Signal Enhancement in AM-FM Interference

5. FUNDING NUMBERS
C — F19628-90-C-0002
PR — 411-2-240

6. AUTHOR(S)
Thomas F. Quatieri, Robert B. Dunn, and Robert J. McAulay

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Lincoln Laboratory, MIT
P.O. Box 73
Lexington, MA 02173-9108

8. PERFORMING ORGANIZATION REPORT NUMBER
TR-993

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
NSMRL
Box 900, Subbase NLON
Groton, CT 06349

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ESC-TR-93-328

11. SUPPLEMENTARY NOTES
None

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

A new approach to interference suppression is developed to enhance the audibility of signals corrupted by amplitude- and frequency-modulated (AM-FM) tonal interference. The suppression algorithm uses a short-time, least squares estimation of the parameters of an AM-FM model of the time-varying tonal interference. The method, developed in a sine-wave analysis/synthesis framework, can be integrated with time and frequency modifications for further signal enhancement. Suppression is applied to single and multitone synthetic and actual AM-FM interference, the latter including man-made signals (e.g., siren interference) as well as naturally occurring signals (e.g., biologic interference). The relative advantages and disadvantages of the sine-wave framework in contrast to a short-time Fourier transform overlap-add framework are described. The enhancement techniques are robust in a large range of environments and can be designed to preserve a random noise background. Finally, it is shown that interference suppression on multichannels prior to beamforming enhances beamformer performance.

14. SUBJECT TERMS
signal enhancement; audibility improvement; AM-FM tonal interference; interference suppression; multitone interference; short-time, least-squared error estimation; beamforming; background preservation; sine-wave analysis/synthesis

15. NUMBER OF PAGES
82

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Same as Report